

Titre: Exploration et traitement multidonnées appliqués à des corridors d'autobus
Title:

Auteur: André Lomone
Author:

Date: 2014

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Lomone, A. (2014). Exploration et traitement multidonnées appliqués à des corridors d'autobus [Mémoire de maîtrise, École Polytechnique de Montréal].
Citation: PolyPublie. <https://publications.polymtl.ca/1357/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/1357/>
PolyPublie URL:

Directeurs de recherche: Robert Chapleau
Advisors:

Programme: Génie civil
Program:

UNIVERSITÉ DE MONTRÉAL

EXPLORATION ET TRAITEMENT MULTIDONNÉES
APPLIQUÉS À DES CORRIDORS D'AUTOBUS

ANDRÉ LOMONE

DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE CIVIL)

FÉVRIER 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

EXPLORATION ET TRAITEMENT MULTIDONNÉES
APPLIQUÉS À DES CORRIDORS D'AUTOBUS

présenté par : LOMONE André

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. TRÉPANIÉ Martin, Ph.D., président

M. CHAPLEAU Robert, Ph.D., membre et directeur de recherche

M. GRONDINES Jocelyn, M.Sc.A., membre

REMERCIEMENTS

L’auteur exprime sa gratitude envers son directeur de recherche, Robert Chapleau, pour son soutien inestimable. Marie-Christine Desharnais, étudiante au doctorat au moment de la rédaction de ce mémoire, et les personnes du groupe MADITUC — Bruno Allard, Daniel Piché, Vincent Dionne et Guillaume Bisailon — sont aussi chaleureusement remerciés pour leur expertise et pour leurs précieux conseils.

Les remerciements vont aussi aux parents de l’auteur et aux organismes suivants pour leur générosité et pour leur soutien financier :

- ministère des Transports du Québec;
- Fonds de recherche du Québec — Nature et technologies;
- département des génies civil, géologique et des mines de l’École Polytechnique de Montréal;
- Association des transports du Canada;
- Fondation de Polytechnique;
- Association canadienne du transport urbain.

À tous, merci beaucoup!

RÉSUMÉ

La collecte passive des informations et ses faibles coûts, par rapport à la grande quantité de données générées, expliquent l'implantation de plus en plus répandue des systèmes de collecte automatique de données. Ces derniers représentent un formidable potentiel pour mieux caractériser la demande et l'offre de transport en commun pour la planification opérationnelle. Toutefois, en plus de leur traitement, l'intégration des données de plusieurs systèmes constitue un important défi.

Cette étude vise à développer des outils et des méthodes pour exploiter ces types de données. La réalisation de cette étude, dans un contexte de recherche et d'analyse exploratoires, se fait dans la perspective d'intégration des différentes sources de données. Ce projet se sert des corridors d'autobus achalandés des boulevards Saint-Michel et Pie-IX du réseau de la Société de transport de Montréal comme exemple d'application des méthodes développées.

D'abord, le réseau planifié est estimé à partir des données GTFS, servant d'information usager : les données des passages aux arrêts sont compressées en séquences de tournées et de voyages. Les voyages non commerciaux et une liste des tracés possibles pour chaque ligne sont déduits à partir de ces données. En classifiant les lignes du réseau, l'estimation des autobus dans le temps dégage les ressources nécessaires au maintien du service. Ensuite, les traces GPS des véhicules et les comptages automatiques (AVL-APC) des passagers permettent d'associer les temps de passage et les mouvements de passagers aux arrêts.

Aussi, les validations de passages (AFC) sont appariées aux arrêts grâce au numéro de véhicule et aux temps de passages obtenus par les traces GPS. Cette opération suppose qu'un usager valide son passage entre l'arrivée de l'autobus à son arrêt et l'arrivée au prochain arrêt. Un modèle de régression linéaire suggère que les compteurs automatiques estiment environ 5 % plus de montées que les validations de passage. Après, une analyse de la ponctualité aux terminus indique, globalement, que les voyages tendent à terminer en avance.

Ensuite, un diagramme distance-temps montre l'évolution des différents voyages dans un corridor. Le diagramme peut confronter les voyages planifiés et réalisés et peut faire ressortir les montées, les descentes et les charges des véhicules. Ce diagramme étant interactif, l'utilisateur peut sélectionner les éléments à afficher et évite de superposer trop d'informations. Enfin, un tableau de bord résume visuellement différents indicateurs et leurs tendances pour chaque arrêt

d'une ligne-tracé : distance interarrêts comme mesurée par le système AVL, ponctualité, durée des haltes, validations, mouvements des usagers, charge, etc.

Cette étude montre le potentiel de l'intégration de données aux sources multiples et constitue quelques pas dans cette direction. Les outils développés permettent de bonifier la compréhension de l'offre et de la demande, autant planifiée que réalisée, ce qui permet de mieux répondre aux préoccupations des planificateurs à propos de l'efficacité et de la performance de leur réseau.

ABSTRACT

Passive information collection and their low costs, considering the large amount of data generated, explain the increasingly more widespread implementation of automatic data collection systems. These data represent a great potential to improve the characterisation of supply and demand of public transit for operational planning. However, in addition to their treatment, the integration of these data from multiple systems is a major challenge.

This study develops tools and systematic methods in order to exploit these types of data. The completion of the study, in an exploratory research and analysis context, has the perspective of integration of multiple-source data. This project uses the crowded Saint-Michel and Pie-IX corridors of Montreal's bus network as examples for the application of the presented methods.

First, the planned network is estimated with GTFS data, used for traveller information: the passage-stop data are compressed by sequences of blocks and runs. Non-commercial runs and line-patterns are inferred from GTFS data. By classifying the network's lines, the evaluation of buses during time reveals the resources needed for the service. Next, arrival and departure time and passenger movements can be associated to stops with GPS tracks and passenger counting data (AVL-APC).

Also, validations (AFC) data are matched with stops by using vehicle number and arrival-departure time, derived from GPS tracks. To perform this operation, it is assumed that a transaction is validated between the vehicle arrival and the arrival to the next stop. A linear regression model suggests that the passenger counters measure about 5 % more boardings than AFC system. Next, a punctuality analysis for terminuses shows that trips tend to finish early.

Then, an interactive distance-time diagram illustrates the evolution of the different trips in a corridor. The diagram can confront planned trips and « actual » trips and can show ridership and most used stops. Because the diagram is interactive, the user can select elements to display and avoids superposing too much information on the figure. Finally, an interactive dashboard visually summarizes many indicators and their trends for each stop of a line-pattern: distance between stops as measured by AVL system, punctuality, halt duration, validations, passenger movements, load, etc.

This study shows the potential to integrate multiple source data and forms a few steps in this direction. These developed tools can enhance the understanding of supply and demand, as planned as realised, which helps to address the planners' concerns about the effectiveness and the performance of their network.

TABLE DES MATIÈRES

REMERCIEMENTS	III
RÉSUMÉ	IV
ABSTRACT	VI
TABLE DES MATIÈRES	VIII
LISTE DES TABLEAUX	XI
LISTE DES FIGURES	XII
LISTE DES SIGLES ET ABRÉVIATIONS	XV
LISTE DES ANNEXES	XVIII
CHAPITRE 1 INTRODUCTION.....	1
1.1 Contexte et problématique.....	1
1.2 Objet d'étude	2
1.3 Contenu et structure	2
CHAPITRE 2 FONDEMENTS ET CONCEPTS.....	5
2.1 Approches de modélisation des données.....	5
2.1.1 Procédure séquentielle classique.....	5
2.1.2 Approche totalement désagrégée	8
2.1.3 Approche orientée objet	10
2.2 Systèmes ADC	11
2.2.1 Multiplication des données : généralités	11
2.2.2 Potentiel des données	11
2.2.3 Défi d'intégration des données	14
2.3 Organisation générale du service	15
CHAPITRE 3 DONNÉES	18

3.1	Données GTFS et estimation du réseau planifié.....	18
3.1.1	Données de la STM.....	18
3.1.2	Détection des lignes-tracés et compression	20
3.1.3	Aperçu du réseau planifié	21
3.2	Systèmes AVL-APC	26
3.2.1	Données brutes.....	27
3.2.2	Détection des arrêts	29
3.2.3	Balancement des mouvements de passagers.....	30
3.3	Système AFC	32
3.3.1	Acquittement du droit de passage	32
3.3.2	Traitements	33
3.3.3	Validations	33
CHAPITRE 4	EXPLORATION ET TRAITEMENT DES DONNÉES.....	36
4.1	Choix des corridors	36
4.2	Représentation de certains indicateurs	39
4.3	Confrontation AFC-APC.....	40
4.3.1	Modèles de régression et appariement aux arrêts	40
4.3.2	Application aux corridors	47
4.4	Ponctualité en début et en fin de voyage	51
4.5	Diagramme distance-temps interactif	56
4.5.1	Construction et description du diagramme	56
4.5.2	Application aux corridors	59
4.6	Tableau de bord	68
4.6.1	Construction et description du tableau de bord	68

4.6.2	Application aux corridors	71
4.7	Discussion.....	85
CHAPITRE 5 CONCLUSION		87
5.1	Résultats	87
5.2	Perspectives de recherche.....	88
BIBLIOGRAPHIE		91
ANNEXES		102

LISTE DES TABLEAUX

Tableau 3-1. Décortication de l'identifiant d'un voyage d'autobus de la STM	19
Tableau 3-2. Classification des lignes d'autobus	23
Tableau 3-3. Estimation des véhicules-heures pour le réseau planifié 2010 (jour de semaine)	24
Tableau 3-4. Estimation des voyages pour le réseau planifié 2010 (jour de semaine)	24
Tableau 3-5. Lignes en fonction du nombre de tracés possibles (réseau 2010, jour de semaine) .	24
Tableau 3-6. Attributs possibles pour chaque entrée du fichier de données du système AVL-APC	28
Tableau 3-7. Estimation des voyages réels associés aux voyages GTFS	30
Tableau 3-8. Différence entre les montées et les descentes « balancées »	32
Tableau 3-9. Validations de passages dans la base de données	33
Tableau 3-10. Validations de passages par mode et par type d'activité.....	34
Tableau 4-1. Corridors ayant le plus de voyages détectés.....	37
Tableau 4-2. Arrêts où les différences entre les données AFC et APC sont les plus grandes (véhicule 22289, ligne 139).....	50

LISTE DES FIGURES

Figure 1-1. Organisation du mémoire	4
Figure 2-1. Procédure séquentielle classique	6
Figure 2-2. Structure informationnelle d'un déplacement totalement désagréé	9
Figure 2-3. Définition schématisée de divers objets et leurs relations dans un réseau d'autobus ..	16
Figure 2-4. Exemple simple d'une tournée d'autobus.....	17
Figure 3-1. Données GTFS de la STM.....	19
Figure 3-2. Compression des stop_times en séquences de voyages et détection des lignes-tracés : exemple avec une tournée fictive	21
Figure 3-3. Estimation des autobus utilisés pour le réseau planifié 2010 (jour de semaine)	25
Figure 3-4. Lignes-tracés 2010, reconstruites par segments à vol d'oiseau entre les arrêts.....	26
Figure 3-5. Système AVL-APC avec un autobus régulier à deux portes	27
Figure 3-6. Aperçu d'une partie de la trace GPS du véhicule 21211	28
Figure 3-7. Estimation des autobus en fonction du temps sur le réseau planifié, en conservant seulement les voyages GTFS associés à un voyage réel.....	30
Figure 3-8. Distribution des véhicules selon la différence montées-descentes	31
Figure 3-9. Validations de titres en fonction du mode et du type d'activité dans le temps.....	34
Figure 4-1. Corridors des lignes 67/467 et 139/439.....	38
Figure 4-2. Corridor Pie-IX (tracés des lignes 439/505)	39
Figure 4-3. Exemples de boîtes à moustaches et centiles utilisés	40
Figure 4-4. Modèle de régression linéaire confrontant les données brutes APC-AFC par véhicule	41
Figure 4-5. Arrêts jalons pour l'appariement des données AFC aux arrêts d'un voyage productif	43
Figure 4-6. Modèle de régression linéaire confrontant les données APC-AFC par véhicule.....	45

Figure 4-7. Modèle de régression linéaire confrontant les données APC-AFC par voyage.....	46
Figure 4-8. Différences entre les données AFC-APC par voyage dans le temps	47
Figure 4-9. Modèle de régression linéaire confrontant les données APC-AFC par voyage, lignes des corridors exemples.....	48
Figure 4-10. Évolution temporelle du cumul des montants, véhicule 22289	49
Figure 4-11. Différences par arrêt entre les données AFC et APC du véhicule 22289.....	50
Figure 4-12. Ponctualité en début et en fin de voyage selon le type de voyage	53
Figure 4-13. Répartition de la différence entre les heures de passage réelles et planifiées en début et en fin de voyage	54
Figure 4-14. Ponctualité des lignes productives en début et en fin de voyage	55
Figure 4-15. Données entrantes et génération du diagramme distance-temps.....	57
Figure 4-16. Extrait de diagramme distance-temps et exemple d'infobulle, corridor Saint-Michel	58
Figure 4-17. Extrait de diagramme distance-temps, ligne 67 sud, voyages GTFS et « réels »	60
Figure 4-18. Diagramme distance-temps, ligne 467 sud	61
Figure 4-19. Diagramme distance-temps, voyage 139_004_016, ligne 139 nord	62
Figure 4-20. Diagramme distance-temps, corridor Saint-Michel, direction sud, entre 08:00 et 09:00	64
Figure 4-21. Diagramme distance-temps, corridor Saint-Michel, direction sud, entre 08:00 et 09:00, avec charges.....	65
Figure 4-22. Diagramme distance-temps, corridor Pie-IX, direction nord, avec charges, montées et descentes.....	67
Figure 4-23. Données entrantes et génération du tableau de bord	68
Figure 4-24. Exemple d'infobulle pour les objets du tableau de bord.....	69
Figure 4-25. Tableau de bord de la ligne 467, direction sud	74
Figure 4-26. Tableau de bord de la ligne 467, direction nord.....	75

Figure 4-27. Tableau de bord de la ligne 67, direction nord.....	76
Figure 4-28. Tableau de bord de la ligne 67, direction sud	77
Figure 4-29. Tableau de bord de la ligne 505, direction nord (tracé 1).....	78
Figure 4-30. Tableau de bord de la ligne 505, direction sud (tracé 2)	79
Figure 4-31. Tableau de bord de la ligne 505, direction sud (tracé 3)	80
Figure 4-32. Tableau de bord de la ligne 139, direction nord.....	81
Figure 4-33. Tableau de bord de la ligne 139, direction sud	82
Figure 4-34. Trace GPS du véhicule 26076.....	83
Figure 4-35. Trace GPS du véhicule 29825	84

LISTE DES SIGLES ET ABRÉVIATIONS

ADC	(~ <i>system</i>) <i>automatic data collection</i> (système ~) (collecte automatique de données)
AFC	(~ <i>system</i>) <i>automatic fare collection</i> (système ~) (perception automatique des titres de transport, des droits de passage)
AMT	Agence métropolitaine de transport
AOT	autorité organisatrice de transport (en commun)
APC	(~ <i>system</i>) <i>automatic passenger counting, counter</i> (système ~) (comptage, compteur automatique de passagers)
AVL	(~ <i>system</i>) <i>automatic vehicle location</i> (système ~) (localisation automatique du véhicule)
BAM	boite à moustaches
CAP	carte à puce
CATI	<i>computer-assisted telephone interview</i> (entrevue téléphonique assistée par ordinateur)
CSS	Cascading Style Sheets (feuille de style en cascade)
D	descente ou descentes
D3	(bibliothèque JavaScript ~) Data-Driven Documents
D3.js	(voir D3)
DABI	<i>driver-assisted bus interview</i> (entrevue dans l'autobus assisté par le chauffeur)
GIS	<i>geographic information system</i> (voir SIG)

GPS	<i>global positioning system</i> (système mondial de localisation)
GTF	(format ~, spécification ~) General Transit Feed (voir GTFS)
GTFS	General Transit Feed Specification (spécification de flux (de données) général de transport en commun)
HLP	haut-le-pied
HTML	HyperText Markup Language (langage de balisage hypertexte)
INIT	Init Innovation in traffic systems AG, ou ses filiales
JS	JavaScript
MADITUC	(groupe (de recherche) ~) Modèle d'analyse désagrégée des itinéraires de transport urbain collectif
M	montée ou montées
MD	montées et descentes
MDDEP	ministère du Développement durable, de l'Environnement et des Parcs du Québec (2005-2012)
MTQ	ministère des Transports du Québec
N/A	non applicable
OD	(enquête ~, matrice ~, paire ~) origine-destination
OQLF	Office québécois de la langue française
SAEIV	système d'aide à l'exploitation et d'information usager
SIG	système d'information géographique
SLR	système léger sur rail
SRB	service rapide par bus
STL	Société de transport de Laval

STM	Société de transport de Montréal
STO	Société de transport de l’Outaouais
SVG	(~ <i>format</i>) Scalable Vector Graphics (format ~) (graphique vectoriel adaptable)
TC	transport en commun
TRANSIMS	(plateforme ~) Transportation Analysis Simulation System
véh	véhicule ou véhicules
véh-h	véhicules-heures

LISTE DES ANNEXES

ANNEXE 1	INFORMATIONS SUR CERTAINES FIGURES	102
ANNEXE 2	DÉFINITIONS D'OBJETS.....	103
ANNEXE 3	ESTIMATION DES VÉHICULES-HEURES DES VOYAGES GTFS ASSOCIÉS À UN VOYAGE RÉEL	105
ANNEXE 4	VARIATIONS AFC-APC PAR ARRÊT (VÉHICULE 22289, LIGNE 139)	106
ANNEXE 5	DÉTAILS SUR LES OUTILS INTERACTIFS (DIAGRAMMES DISTANCE- TEMPS ET TABLEAUX DE BORD).....	107
ANNEXE 6	FICHIERS ENTRANTS POUR LE DIAGRAMME DISTANCE-TEMPS..	108
ANNEXE 7	PARAMÉTRAGE DU DIAGRAMME DISTANCE-TEMPS, LIGNES 67 ET 467.....	110
ANNEXE 8	FICHER ENTRANT POUR LE TABLEAU DE BORD	111
ANNEXE 9	PARAMÉTRAGE DU TABLEAU DE BORD, LIGNE 467 SUD.....	112

CHAPITRE 1 INTRODUCTION

1.1 Contexte et problématique

Les indices et les preuves s'accumulent : l'augmentation des émissions de gaz à effet de serre par les activités humaines contribue aux changements climatiques, partout sur la planète. Les différents gouvernements se voient maintenant confrontés à ce problème complexe et doivent en tenir compte dans leurs différentes politiques.

Par exemple, le gouvernement québécois a élaboré un *Plan d'action sur les changements climatiques* couvrant 2006 à 2012 (ministère du Développement durable, de l'Environnement et des Parcs du Québec [MDDEP], 2006). Puisque le secteur du transport routier génère la plus grande part des émissions québécoises de gaz à effet de serre, soit environ 33.3 % en 2010 (MDDEP, 2013), le plan inclut notamment un *Programme d'amélioration des services en transport en commun* (ministère des Transports du Québec [MTQ], 2012). Le plan d'action a été renouvelé pour la période 2013-2020 et priorise, entre autres, la promotion du transport collectif par l'amélioration de l'offre de service (MDDEP, 2012).

Aussi, le ministère des Transports du Québec (MTQ, 2009) accorde une importance particulière au transport collectif dans son *Plan stratégique 2008-2012* en adoptant plusieurs axes d'intervention : efficacité et intégration des systèmes de transport, diversification des modes, réduction des gaz à effet de serre, renouvellement des équipements de transport en commun, etc.

La stratégie gouvernementale québécoise et les orientations stratégiques du MTQ sont des exemples qui montrent que le transport collectif est vu de plus en plus comme un élément pouvant être important dans la poursuite du développement durable.

Depuis plusieurs années, diverses autorités organisatrices de transport en commun (AOT) commencent à installer et à expérimenter plusieurs technologies émergentes afin d'améliorer leur productivité. Ces technologies sont intégrées à des systèmes de collecte automatique de données (ADC, *automatic data collection*), qui, comme le nom le présage, enregistrent de l'information sans intervention de l'utilisateur ou presque. Cela inclut souvent un système AVL (*automatic vehicle location*) utilisant les récepteurs GPS pour localiser le véhicule, des compteurs de passagers du

système APC (*automatic passenger counting*) et un système AFC (*automatic fare collection*) qui est relié à la boîte de perception de l'autobus, qui peut lire notamment les cartes à puce (CAP).

Les données de ces systèmes représentent un grand potentiel pour la planification du transport collectif et pour l'amélioration de sa performance et de sa gestion, bien qu'il subsiste encore quelques défis quant à la détection et à la correction des erreurs (Furth, Hemily, Muller, & Strathman, 2006; Utsunomiya, Attanucci, & Wilson, 2006). Pris séparément, tous ces équipements génèrent une quantité très importante de données imparfaites, mais, combinées et intégrées, ces données forment une redondance permettant de s'approcher de la « vraie » information (Chapleau & Allard, 2010), permettant ainsi de mieux faire ressortir les tendances et les idées.

1.2 Objet d'étude

Ce projet de recherche s'inscrit dans une démarche d'intégration de ces types de données. Cet effort de recherche et d'analyse exploratoires inclut le développement d'outils et de méthodes permettant d'exploiter ces données et pouvant être utiles à la planification opérationnelle. Les outils développés devraient ainsi permettre de confronter le service planifié et rendu, de mieux connaître l'offre et la demande et d'évaluer et d'améliorer la performance du réseau.

Ce projet compte appliquer les méthodes présentées sur deux corridors importants du réseau de la STM (Société de transport de Montréal), qui servent alors d'exemples. Les données incluent, d'une part, des données GTFS (General Transit Feed Specification), destinées à de l'information voyageur, et d'autre part, des données collectées par les instruments de plusieurs autobus de la STM.

1.3 Contenu et structure

La figure 1-1 présente la structure et l'organisation du document sous forme de diagramme. D'abord, les fondements et concepts sur lesquels repose cette recherche sont abordés (chapitre 2) pour mettre en relief les enjeux reliés : approches de modélisation des données, potentiel et défis des données des systèmes ADC et organisation générale du service. Le chapitre 3 décrit les données GTFS, avec lesquelles le réseau planifié d'autobus est estimé, et celles des systèmes AVL-APC et AFC.

L'exploration et le traitement de ces données est l'objet principal du chapitre 4. Il y est question d'appariement des données AFC aux arrêts, de confrontation des données sur les montants entre les systèmes AFC et APC et du développement d'outils de visualisation et d'analyse d'indicateurs, comme la ponctualité, la charge des autobus ou la durée des arrêts. Avant de conclure, une discussion est abordée, en exposant notamment les limites du projet. Enfin, un résumé des réalisations et les perspectives de recherche sont présentés en conclusion (chapitre 5).

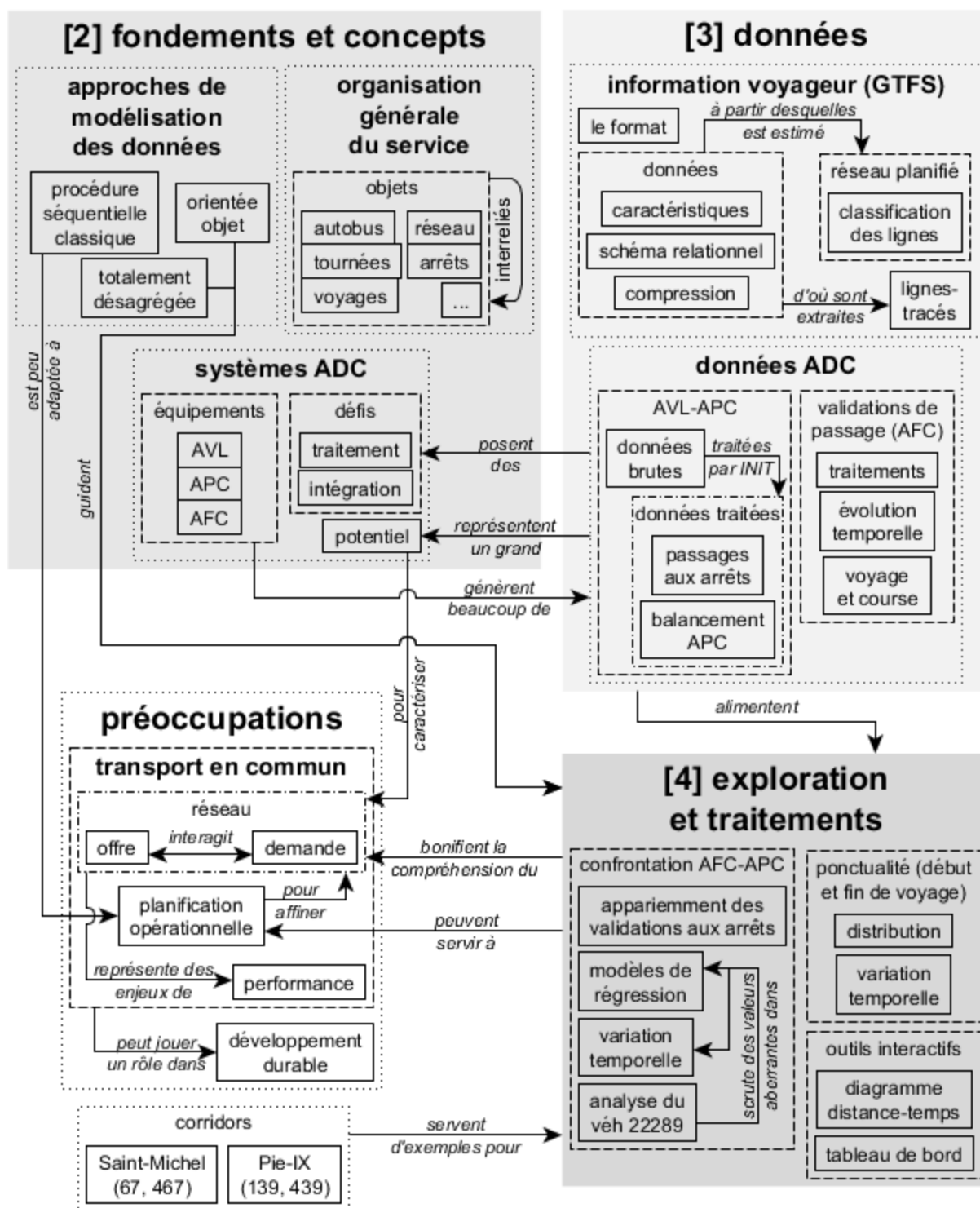


Figure 1-1. Organisation du mémoire

CHAPITRE 2 FONDEMENTS ET CONCEPTS

Ce chapitre établit le contexte dans lequel s'inscrit ce mémoire. D'abord, les approches de modélisation des données sont résumées : rappel et critique de la procédure séquentielle classique; approche totalement désagrégée et approche orientée objet. Le potentiel et les défis qui accompagnent les données des systèmes ADC sont exposés. Enfin, la description de l'organisation générale d'un service d'autobus vient clore ce chapitre.

2.1 Approches de modélisation des données

En plus des comptages à bord, la procédure séquentielle classique a déjà été répandue dans la modélisation des déplacements (Bates, 2008). Or, en plus des nombreuses critiques (résumées à la section 2.1.1), cette procédure est plus adaptée dans un contexte global (Thévenin, 2001). Pour une échelle locale et pour une caractérisation plus fine de la demande et de l'offre en transport, les approches totalement désagrégées et orientée objet conviennent mieux et sont utilisées dans ce projet (voir section 2.1.2 et 2.1.3 respectivement).

2.1.1 Procédure séquentielle classique

Beimborn, Kennedy et Schaefer (1996); Domencich et McFadden (1975); McNally (2008); Ortúzar et Willumsem (2011) et bien d'autres font partie des nombreux auteurs qui ont résumé et documenté cette procédure. La figure 2-1 peint un portrait général de la procédure et les divers modèles entourant ses étapes.

La procédure requiert d'abord un découpage zonal du territoire d'étude. Chaque zone est ainsi décrite par des attributs moyennés à partir de ses habitants et par son centroïde. Ensuite, l'étape de la génération permet de prédire les déplacements produits et attirés pour chaque zone, en utilisant souvent des régressions linéaires multiples. À l'étape de distribution, ces déplacements sont répartis par modèle gravitaire dans la matrice origine-destination (OD) pour former les paires OD entre les zones. Après, le mode de chaque déplacement est estimé en utilisant souvent un modèle probabiliste et une fonction d'utilité. Enfin, les déplacements sont affectés sur le réseau, souvent avec un processus itératif pour équilibrer les flots et pour tenir compte de la capacité des liens.

Certains modèles peuvent tenter une rétroaction en prédisant l'évolution de l'utilisation du sol et du système de transports. D'autres itérations sur la procédure sont alors exécutées.

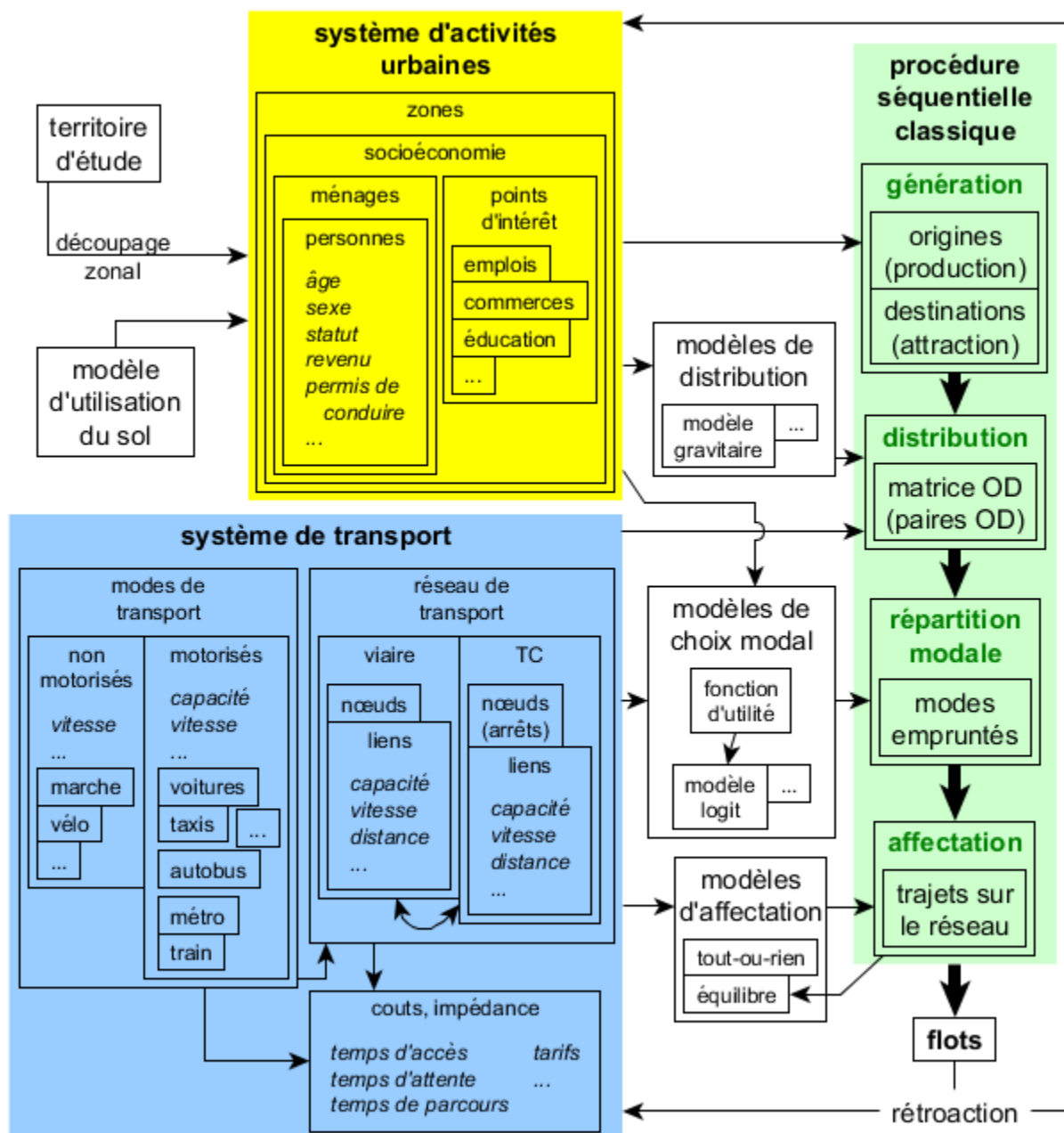


Figure 2-1. Procédure séquentielle classique

Même si les technologies de l'information ont beaucoup évolué depuis les années 1950, la procédure séquentielle classique a longtemps été répandue (Bates, 2008), malgré les discrédances observées, souvent énormes, avec la réalité (Jones, Dix, Clarke, & Heggie, 1983), et est restée

inchangée tout en restant inflexible et inexacte à cause de ses hypothèses (Guiliano, 1985, cité dans Atkins, 1986). De plus, à l'époque, la procédure a déjà souvent reposé sur des données dont les procédures de collecte et dont les mises à jour n'étaient pas une grande préoccupation (Stopher, 1983).

Plusieurs auteurs, dont Atkins (1986); Beimborn et al. (1996); Domencich et McFadden (1975); Goodwin (1984); Manheim (1979), ont critiqué cette méthode en énumérant ses limites, dont les principales sont mises en relief ci-dessous :

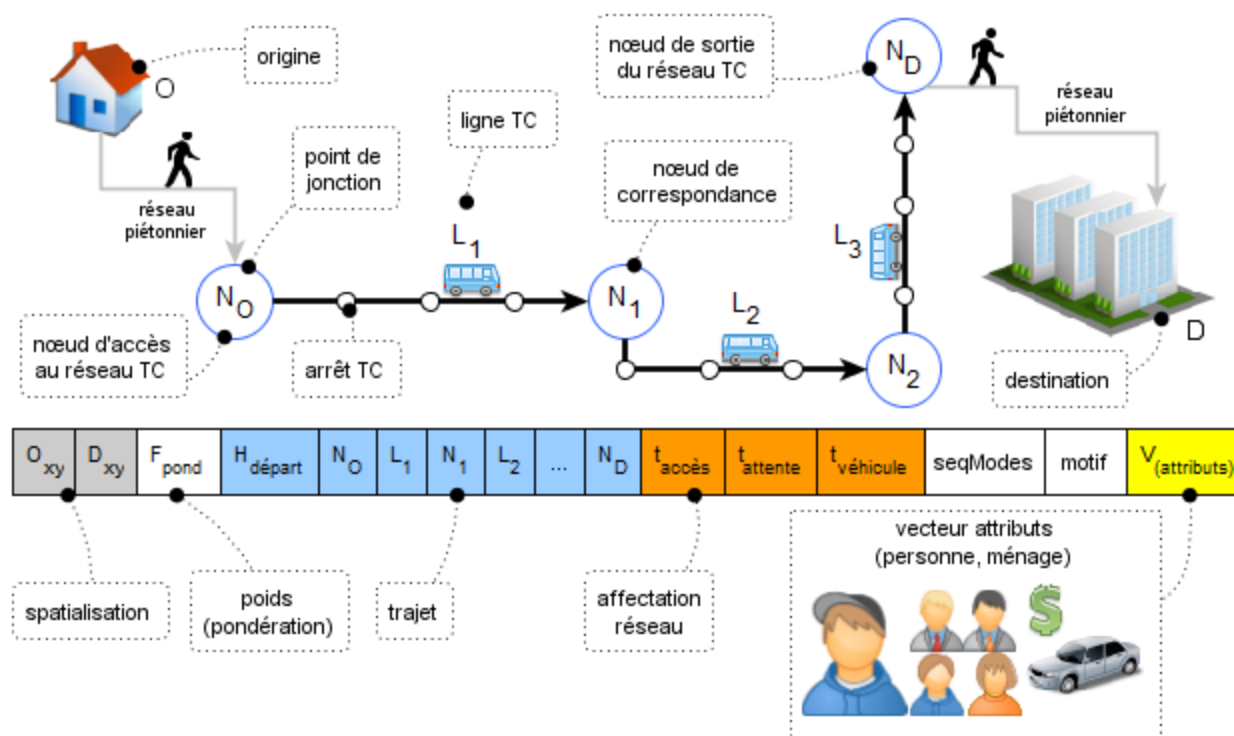
- découpage du territoire — Le recours à l'agrégation par zones implique nécessairement que les comportements sont uniformes pour les personnes d'une zone, limitant l'étude de la variabilité individuelle des déplacements. Cela impose aussi le fait que les déplacements débutent et terminent à un centroïde, ce qui est plutôt irréaliste. Aussi, la délimitation des zones peut modifier fortement les résultats;
- caractère statique et rétroaction — Le caractère séquentiel de la procédure rend la rétroaction difficile. Par exemple, l'étape de distribution nécessite de connaître les temps de parcours qui ne seront véritablement connus et précisés qu'à l'affectation des déplacements sur le réseau;
- flots et comportements — Alors que l'objectif de la procédure est de modéliser les flots sur le réseau, même si elle ne permet pas de les estimer directement, d'autres modèles et approches permettent d'étudier plus finement le comportement individuel des usagers et d'évaluer les infrastructures de transport;
- population synthétique — Le recours à une matrice OD implique que les déplacements générés sont synthétiques et ne révèle donc pas les particularités individuelles des déplacements;
- biais systématiques — S'il est vrai que les comportements de déplacements sont influencés par l'âge, par l'inertie et par les habitudes, alors les biais systématiques engendrés par ces facteurs ne sont pas détectés par les tests statistiques utilisés par les procédures d'estimation;
- l'œuf ou la poule? — Les modèles de génération de déplacements incluent des facteurs qui sont eux-mêmes dépendants des déplacements réalisés. Par exemple, les déplacements attirés pour motif magasinage sont fonction des emplois des commerces. Or, ces mêmes emplois sont influencés par le nombre de clients des commerces;

- influence du système de transport sur la génération — L'étape de génération se fait souvent sans tenir compte du service de transport qui, pourtant, peut fortement influencer le choix de la destination et du mode de transport emprunté;
- estimation biaisée de la matrice OD — Les couts de déplacements utilisés dans la distribution des déplacements pour obtenir les paires OD sont souvent basés uniquement sur le temps ou la distance de parcours en automobile seulement, même s'il est tout de même possible de considérer séparément les modes à l'étape de distribution. Et ces couts ne tiennent pas compte des effets de congestion. Aussi, les facteurs socioéconomiques n'entrent pas dans l'estimation de la matrice. Les déplacements entre des quartiers riches et des centres d'emplois à revenu peu élevé pourraient être surestimés par exemple;
- évaluation du choix modal — Les principales variables du modèle de choix modal restent le temps, ou la distance, et le cout de déplacement. D'autres variables plus qualitatives tout aussi susceptibles d'influencer le choix du mode, comme le confort et la sécurité, sont difficiles à intégrer dans le modèle. De plus, les hypothèses posées pour la fonction d'utilité influencent fortement le modèle généré (Bierlaire, 1998).

2.1.2 Approche totalement désagrégée

En utilisant des modèles de choix modal, Domencich et McFadden (1975) sont parmi les premiers à utiliser de larges échantillons d'observations individuelles. Les données utilisées sont « désagrégées », dans le sens où les individus ne reçoivent pas les attributs moyens de leur zone. Cela permet de mieux observer la variabilité du comportement de mobilité à travers les individus.

L'approche « totalement » désagrégée a fait son apparition plus tard, dans les années 1980 à travers le logiciel MADITUC (modèle d'analyse désagrégée des itinéraires de transport urbain collectif), développé par le groupe de recherche du même nom (Chapleau, 1986; Chapleau, Allard, & Sarrazin-Canova, 1982). Comme le montre la figure 2-2, l'approche se base sur la caractérisation, ou connaissance, fine des déplacements de chaque individu : origine, destination, correspondances, point d'entrée et de sortie du réseau, lignes de transport en commun, temps d'accès, temps d'attente, temps en véhicule, séquence de modes, motif, etc.



(figure reprise de Chapleau, 1992) (voir annexe 1 pour références des icônes)

Figure 2-2. Structure informationnelle d'un déplacement totalement désagrégé

Cette approche se distingue de ses prédécesseurs par l'utilisation des concepts principaux suivants (Trépanier, Chapleau, Allard, & Morency, 2004; Trépanier, Chapleau, & Morency, 2008) :

- traitement des données sur les déplacements individuels sur tout le processus d'analyse, conservant tous les attributs et caractéristiques des déplacements — temps, motif, modes, itinéraires;
- utilisation de coordonnées cartésiennes — ou de monuments ou de places connues — comme système spatial de référence pour les objets spatialisés du système — origine, destination, résidences, stations, arrêts, etc.

Au lieu de diluer, ou de dissoudre, l'information en l'agréant dans une matrice OD, l'approche totalement désagrégée augmente le niveau de résolution en considérant les données sur les ménages, sur les personnes et sur les déplacements (Trépanier et al., 2008). Par la suite, l'agrégation ultérieure des résultats reste possible et laisse la possibilité d'un découpage territorial mieux adapté à l'objet d'étude (Thévenin, 2001). L'approche désagrégée augmente le nombre de

dimensions distinguées par le système d'information, améliorant ainsi les possibilités d'analyse (O'Donnell & Smith David, 2000). L'objectif ultime de cette approche est donc de créer des modèles pouvant reproduire le plus fidèlement possible les dimensions spatiotemporelle, socioéconomique et démographique de la réalité urbaine observée en exploitant la meilleure information disponible (Morency, 2004).

2.1.3 Approche orientée objet

Le principe méthodologique de l'approche orientée objet — ou « orientée-objet » ou « orientée vers l'objet » ou « par objet » (Office québécois de la langue française [OQLF], 2008) —, venant essentiellement du milieu de la programmation et du stockage de données, peut se résumer à créer, à caractériser et à mettre en relation des objets pour en constituer un système (Trépanier, 1999). Cette approche permet d'identifier rapidement les objets, leurs propriétés et leurs méthodes impliqués dans le système de transport (Chapleau, Trépanier, & Allard, 1998, cité dans Trépanier et al., 2008), permettant ainsi une meilleure compréhension de l'usage spécifique de chaque information (Chapleau, Trépanier, & Allard, 1998, cité dans Desharnais, 2009). L'approche orientée objet s'avère une continuité de l'approche désagrégée en fournissant un cadre pour la structure des données et pour les procédures de leur traitement (Trépanier, Barj, Dufour, & Poilpré, 2004).

Cette approche fait intervenir le concept de classe, une espèce de moule ou de modèle, servant à créer les objets (instances), décrits par leurs propriétés (attributs, composante statique) et par leurs méthodes (actions, composante dynamique) (Trépanier & Chapleau, 2001). Avec le principe d'héritage, les classes peuvent échanger leurs caractéristiques avec leurs sous-classes, formant ainsi une famille où les méthodes sont généralisées (Aubert et Dix-Neuf, 1991, cité dans Trépanier & Chapleau, 2001). L'approche orientée objet peut venir ainsi compléter l'approche totalement désagrégée en catégorisant et en définissant chaque objet impliqué dans les déplacements.

2.2 Systèmes ADC

2.2.1 Multiplication des données : généralités

L'amélioration de la performance des équipements informatiques et électroniques a facilité la collecte et le traitement de grandes quantités de données, autrefois peu abondantes — voire rares — et fragmentées. Par exemple, le Grand collisionneur de hadrons, l'accélérateur de particules construit entre autres pour vérifier l'existence du boson de Higgs, a produit 13 pétaoctets de données (soit 13×10^{15} octets) en 2010 (Brumfield, 2011). Aussi, le décodage du génome humain — contenant environ trois milliards de paires de bases —, terminé la première fois en 2003, a nécessité un travail colossal de plus de dix années, alors que cette opération peut désormais s'effectuer en une semaine (The Economist, 2010a).

Certains désignent ce phénomène par l'émergence des données massives ou volumineuses (*big data*), c'est-à-dire dont la quantité excède les capacités des systèmes traditionnels de base de données (Dumbill, 2012), même si cette notion de quantité reste encore très relative. La multiplication et la facilité d'accès des données renferment un tel potentiel que la spéculation et les hypothèses pourraient être mises de côté : les machines guident maintenant le chercheur vers les modèles et les tendances (Graham, 2012). Toutefois, de bien grands défis s'ajoutent à ces données : comment les organiser, leur donner un sens, et comment former la prochaine génération à propos de cette « révolution industrielle des données »? (The Economist, 2010a) Il ne faut pas oublier non plus le rôle des spécialistes pour les replacer dans leur contexte : quelles idées semblent — ne pas — véhiculer ces données? (Graham, 2012)

2.2.2 Potentiel des données

Le principal défi de la recherche dans le domaine du transport est représenté par les données, autant sur le plan de la collecte que du traitement. Les enquêtes ménages à grande échelle (ou enquêtes OD) peuvent fournir un grand échantillon de déplacements associés à des attributs socioéconomiques sur les personnes. La méthode d'enquête utilise souvent la CATI (*computer-assisted telephone interview*, entrevue téléphonique assistée par ordinateur) : les intervieweurs sont assistés en temps réel par un questionnaire interactif et par une base de données géoréférencées (grands générateurs de déplacements, infrastructures de transport, réseau de transport en commun, etc.) (Chu, Chapleau, & Trépanier, 2009). Comme le résumait Chu et

al. (2009), les enquêtes OD ne sont toutefois pas adaptées au contexte de planification opérationnelle :

- incompatibilité temporelle — Ces enquêtes ne sont effectuées qu'une fois tous les quelques années (tous les cinq ans dans le cas de la grande région de Montréal (Agence métropolitaine de transport de Montréal, 2010)), alors que le service a besoin de multiples ajustements dans l'année;
- échantillon insuffisant — L'échantillon ne vise pas uniquement les déplacements en transport en commun et intègre d'autres modes comme l'automobile;
- résolution inadéquate — La planification opérationnelle exige une résolution spatiotemporelle des données que ne peut pas fournir ces enquêtes à cause de leur méthodologie et des répondants;
- coûts importants — Les enquêtes OD sont coûteuses par rapport aux données qu'elles peuvent fournir. Les enquêtes spécialisées, comme les enquêtes à bord, peuvent résoudre certains des problèmes mentionnés précédemment, mais elles requièrent tout de même un travail et des dépenses élevés.

Or, les systèmes automatiques de collecte de données (ADC pour *automatic data collection*) deviennent de plus en plus répandus en transport collectif. En effet, ces systèmes deviennent accessibles financièrement tout en offrant de riches informations (Zhao, Rahbee, & Wilson, 2007). Dans le cas des autobus, ces systèmes incluent souvent des récepteurs GPS pour le système automatique de localisation de véhicule (AVL pour *automatic vehicle location*), des compteurs automatiques des passagers (APC pour *automatic passenger counter*) et des boîtes de perception des titres de transport (AFC pour *automatic fare collection*) utilisant la technologie des cartes à puce (CAP).

La collecte automatique des données par ces systèmes peut être considérée comme une forme d'enquête passive : par opposition à une enquête origine-destination (OD) traditionnelle ou classique où les personnes interrogées doivent elles-mêmes fournir de l'information, les systèmes ADC collectent de l'information sans que l'utilisateur n'ait à fournir d'effort particulier (Chu, 2010). Ce mode d'acquisition de l'information, aussi appelé DABI (*driver-assisted bus interview* ou

entrevue dans l'autobus assistée par le chauffeur), présente plusieurs propriétés intéressantes : haut niveau de résolution — autant spatiale que temporelle —, grande étendue et couverture des échantillons, nature désagrégée de l'information, intégration avec les données opérationnelles, etc. (Chu et al., 2009)

L'introduction des systèmes ADC s'explique, selon Chu (2010) et Strathman et al. (2008), par une réorientation des préoccupations des opérateurs de transport en commun vers la performance et vers le client et par l'abondance des données apportée par ces systèmes, alors que ces informations étaient difficiles à obtenir à cause de l'emploi de méthodes manuelles et des coûts associés relativement élevés. Et Zhao (2004) ajoute que cette nouvelle stratégie de collecte de données, combinée à des analyses adaptées, aide à mieux supporter le processus de décision des autorités de transport.

Le potentiel du traitement de ces données archivées est mis de l'avant dans la récente littérature. En particulier, le livre *Schedule-based modeling of transportation networks: theory and applications*, de Wilson et Nuzzolo (2009), y consacre plusieurs chapitres. Notamment, Wilson, Zhao et Rahbee (2009) montrent dans quels domaines une autorité de transport peut profiter de l'utilisation de ces données : planification opérationnelle, gestion de la performance, information usager et planification du service.

D'ailleurs, Furth et al. (2006) confirment l'importance des données des systèmes ADC en tant qu'opportunité de caractérisation fine de l'offre et de la demande, en exposant quelques exemples d'analyses utiles : analyses temporelles et opérationnelles (ponctualité, ajustement des temps prévus de parcours, régularité du service, temps d'attente des usagers, etc.), analyses de la demande (profils de charge, montées, descentes, etc.), vérification de la base de données géographiques (comme l'emplacement des arrêts), analyses couplées avec les SIG (systèmes d'information géographique), etc.

De plus, El-Geneidy et al. montrent plusieurs applications dans la planification opérationnelle à partir de données AVL-APC sur le corridor d'autobus Saint-Michel. L'évaluation des temps de parcours des autobus a conduit à une proposition d'un service à arrêts limités (Tétreault & El-Geneidy, 2009), dont la STM s'est inspiré pour implanter une ligne express (Tétreault, 2009). Ce nouveau service sur ce corridor a ensuite fait l'objet d'un suivi en étudiant

les temps de parcours (El-Geneidy & Surprenant-Legault, 2010; Surprenant-Legault & El-Geneidy, 2011) et ses impacts sur la perception du service par les usagers (Diab & El-Geneidy, 2012).

2.2.3 Défi d'intégration des données

Toutefois, pris séparément, chaque système ADC ne peut rendre qu'une information partielle et imparfaite : le taux de récupération de ces données se situe bien souvent entre 25 % et 75 % (Chapleau & Allard, 2010). En effet, la présence d'une grande portion de données incorrectes nécessite des stratégies pour surmonter ces difficultés et pour réaliser le plein potentiel des systèmes ADC (Furth et al., 2006). Alors que Slavin, Rabinowicz, Brandon, Flammia et Freimer (2009) s'intéressent à l'avènement de grandes quantités de données AFC pour mieux comprendre les choix de parcours sur de grands réseaux de transport en commun, ces auteurs exposent l'importance des données périphériques : horaires, réseau SIG et informations de voyage du journal de bord reliées aux validations de passage.

Notamment, les systèmes AFC font l'objet de plusieurs études pour leur potentiel dans la dérivation des caractéristiques de déplacements (Jang, 2010; Park, Kim, & Lim, 2008; Seaborn, Attanucci, & Wilson, 2009) et de la performance du service (Trépanier, Morency, & Agard, 2009; Uniman, Attanucci, Mishalani, & Wilson, 2010). D'ailleurs, Utsunomiya et al. (2006) exposent certains aspects intéressants du potentiel de ce type de données, mais indiquent qu'il reste encore beaucoup de travail pour la détection et la correction des erreurs dans les données. Avec des données AFC, Chu et al. insistent sur le développement de procédures de correction et d'enrichissement des données permettant d'améliorer la cohérence des informations désagrégées des déplacements : âge, motif, origine, destination, arrêt de montée et de descente, générateurs de déplacements, etc. (Chu, 2010; Chu & Chapleau, 2007, 2008, 2010; Chu et al., 2009)

La démarche utilisée dans ce mémoire s'inscrit dans celle utilisée par Chapleau et Allard (2010), qui ont développé des indicateurs de performance et de demande en intégrant des données GPS, APC et AFC. Ils proposent un cadre de modélisation qui met l'accent sur la définition des objets, sur la nécessité de s'attaquer au manque de cohérence des données, sur la communication de l'information par des techniques de visualisation appropriées et sur le développement de procédures de correction et d'enrichissement des données (Chapleau & Allard, 2010).

Aussi, Faivre d'Arcier (2012) combine plusieurs sources de données des réseaux de Montpellier et de Toulouse en France — systèmes APC et AFC, enquêtes OD et recensement — pour guider sa réflexion sur les dimensions commerciale et productive de l'efficacité des réseaux de transport en commun. Toutefois, cette étude soulève un besoin de standardisation de certaines données d'enquêtes et de géolocalisation des informations relatives à l'offre et à la demande de transport et au territoire d'étude (Faivre d'Arcier, 2012).

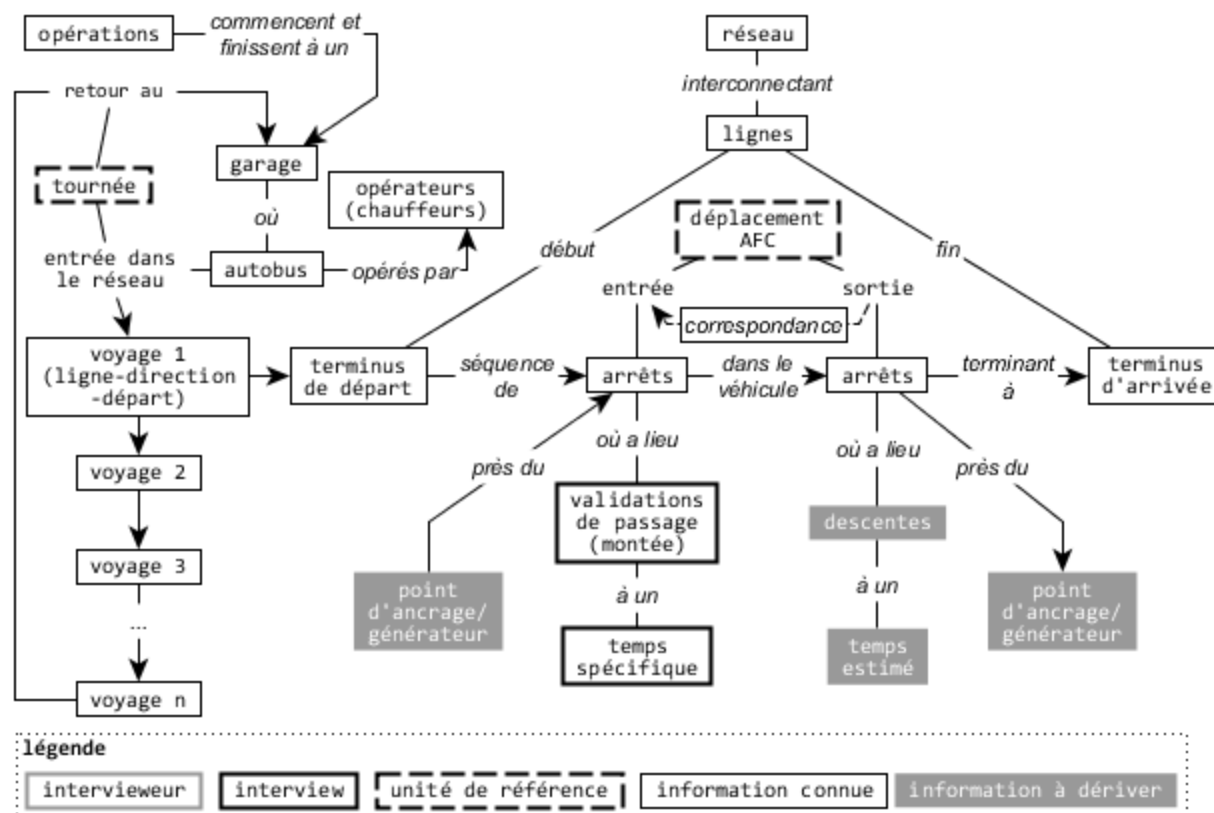
Bref, les données des systèmes ADC peuvent apporter un éclairage intéressant sur le comportement des usagers du transport en commun et sur son réseau. Or, pour atteindre leur plein potentiel, il faut relever les défis de traitement et d'intégration de ces données venant de systèmes différents.

2.3 Organisation générale du service

La figure 2-3 illustre différents objets et leurs relations dans un réseau d'autobus comme celui de la STM (l'annexe 2 résume les définitions des objets cités dans ce mémoire).

Du point de vue opérationnel, les opérations sont divisées en tournées (*vehicle blocks*). Chaque tournée est constituée d'une séquence de voyages. Un voyage peut desservir une ligne-direction avec un départ à une heure déterminée : ce voyage est alors en service commercial, ou productif. Le véhicule suit alors une séquence d'arrêts entre les terminus de départ et d'arrivée de la ligne-direction. D'ailleurs, pour éviter une confusion entre l'arrêt comme objet — l'endroit prévu où l'autobus doit s'immobiliser — et l'arrêt comme acte de s'immobiliser, ce dernier sens est désigné par « halte » dans ce mémoire.

Les usagers entrent et correspondent sur le réseau et en sortent habituellement par les arrêts. Certains de ces arrêts peuvent servir de points de contrôle pour les opérateurs afin de suivre et d'ajuster le service.



(figure reprise et traduite de Chu et al., 2009)

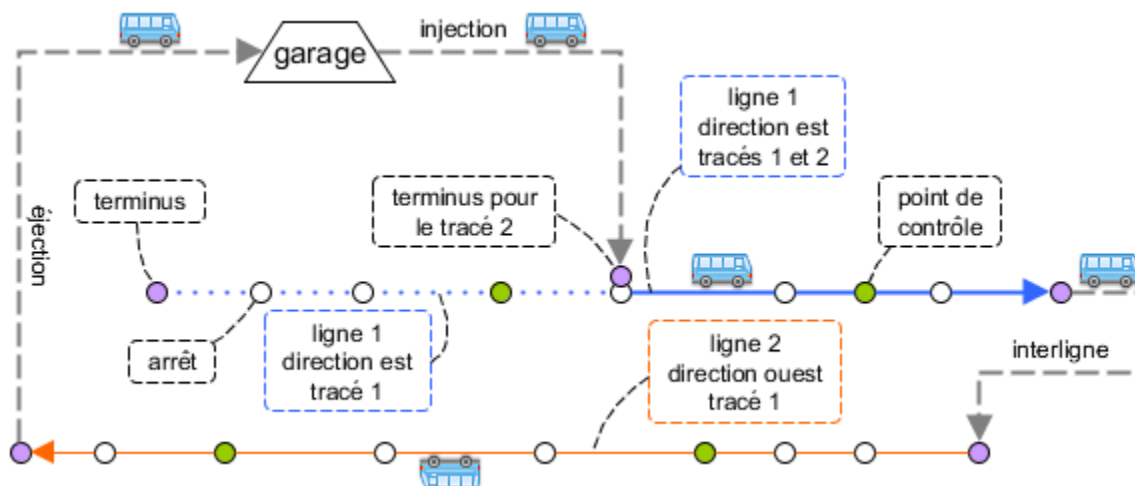
Figure 2-3. Définition schématisée de divers objets et leurs relations dans un réseau d'autobus

Lorsque l'autobus — ou plutôt le chauffeur — a terminé son voyage sur une ligne et attend pour son prochain voyage, au même arrêt ou à proximité, le chauffeur est en battement. Les temps de battement et d'interligne peuvent inclure le temps de repos du chauffeur ou le temps de recouvrement en cas de retard accumulé au voyage précédent.

Un voyage d'une tournée peut aussi servir de haut-le-pied (HLP) entre un garage et le point d'injection dans le réseau, entre le point d'éjection et un garage ou entre deux voyages commerciaux (interligne) (voir figure 2-4). Il n'y a alors généralement que deux arrêts pour ce type de voyage (non commercial ou improductif), soit les terminus.

Une tournée est généralement répétée tous les jours de semaine – d'autres tournées sont planifiées pour le service en fin de semaine et des jours fériés –, mais elle peut être effectuée par des véhicules différents.

Du point de vue du réseau, chaque validation de passage est logiquement liée à un voyage et à un point de montée spécifique. Dans un réseau de lignes préétablies, les montées se produisent le long d'une ligne avec des arrêts dans un ordre préétabli.



(voir annexe 1 pour références des icônes)

Figure 2-4. Exemple simple d'une tournée d'autobus

La figure 2-4 montre un exemple simple d'une tournée d'autobus où le véhicule est injecté au « milieu » d'une ligne. Chaque ligne peut être effectuée dans deux directions différentes habituellement, par exemple dans un axe est-ouest, d'où vient le concept de ligne-direction. Pour chaque direction, il peut y avoir des parcours différents. Cela est le cas, par exemple, lorsqu'une ligne a une séquence d'arrêts différente le soir de celle du matin pour ajuster le service à l'achalandage. Le concept, géométrique, de ligne-tracé (ou tracé, *line-pattern*) intervient alors. Une ligne-tracé est donc une séquence d'arrêts desservis par l'autobus pour une ligne-direction donnée. Lorsqu'une ligne-direction n'a qu'un seul tracé possible, les deux notions — ligne-direction et ligne-tracé — peuvent se confondre.

CHAPITRE 3 DONNÉES

Ce chapitre vise à décrire les données sur lesquelles repose le chapitre suivant. Dans un premier temps, des données GTFS sont mises à profit dans l'estimation du réseau planifié de la STM. Le format de fichiers est brièvement expliqué d'abord. Une note sur la compression des données en lignes-tracés est présentée avant de poursuivre avec l'estimation des ressources impliquées dans le réseau planifié, dans laquelle une classification des lignes est aussi proposée.

Dans un deuxième temps, les données des systèmes collectant automatiquement des données à bord des autobus de la STM sont abordées. Les données des compteurs de passagers et des récepteurs GPS sont vues d'abord. Les données brutes sont brièvement décrites avant de poursuivre avec les données traitées. Par la suite, les données du système de perception automatique des titres sont montrées.

3.1 Données GTFS et estimation du réseau planifié

Développé initialement par Google et par le Tri-County Metropolitan Transportation District of Oregon (Roth, 2010), le GTFS (General Transit Feed Specification, qui pourrait être traduit par « spécification de flux (de données) général de transport en commun ») définit un format commun pour les horaires et pour les informations géographiques associées (Google, 2012a). Le format, dont les spécifications sont décrites sur le site de Google (2012b), consiste en un dossier compressé (.zip) contenant les fichiers texte (.txt) décrivant les arrêts, les lignes, les voyages, les horaires, etc. (Google, 2012a)

Les organismes de transport public peuvent se servir du format GTF pour fournir à Google Maps ou à d'autres services de Google les informations sur leur réseau (Google, 2012a). Ce format permet aussi à ces organismes de publier leurs données qui seront utilisées de manière interopérable par les applications que les développeurs pourront créer (Google, 2012a).

3.1.1 Données de la STM

L'allure des fichiers correspondant au premier trimestre de 2010, pour correspondre aux données AVL-APC et AVL qui datent du mercredi 17 mars 2010, est dessinée à la figure 3-1.

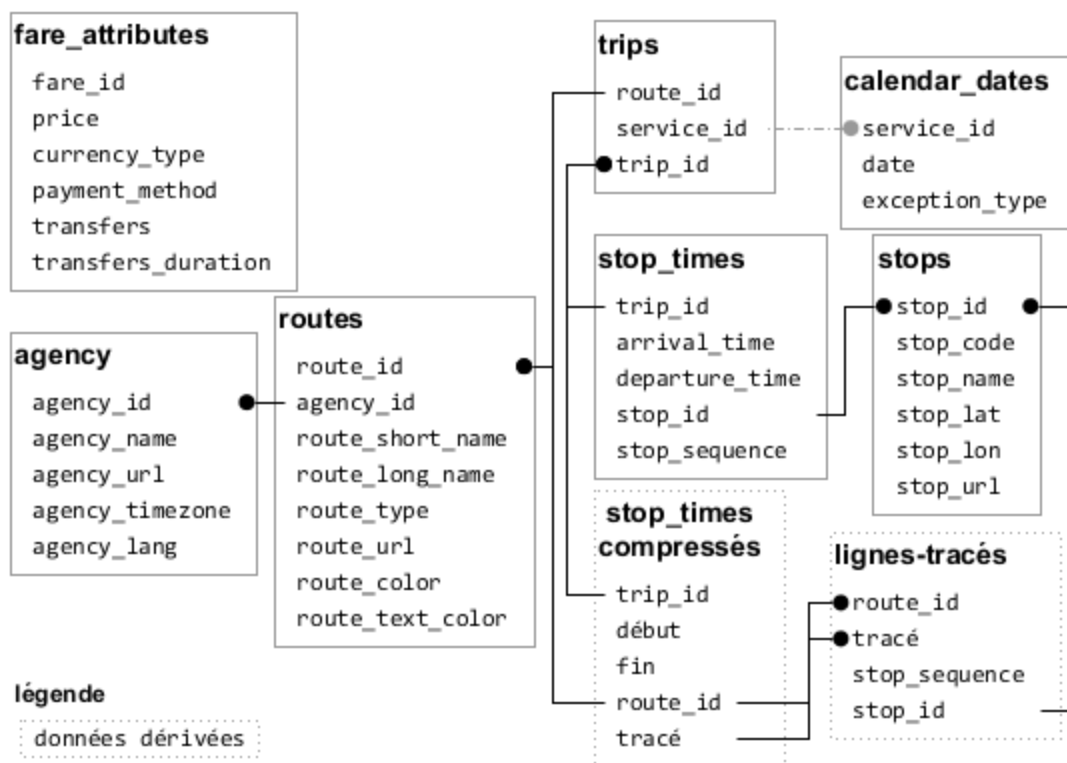


Figure 3-1. Données GTFS de la STM

L'information contenue dans les fichiers de la STM présente quelques caractéristiques qui leur sont propres. D'abord, chaque voyage (*trip*) de la STM a un identifiant permettant de connaître sa tournée correspondante, comme le montre le tableau 3-1. L'identifiant met ainsi en lumière la séquence des voyages de la tournée.

Tableau 3-1. Décortication de l'identifiant d'un voyage d'autobus de la STM

catégorie	élément	exemple
identifiant	trip_id	10J_S_10_1_2_10
horaire	liste (année, mois)	10J (janvier 2010)
	journée de service	S (semaine)
tournée (tâche)	ligne dominante	10
	numéro de tournée	1
voyage	numéro de voyage	2
service	ligne d'autobus	10

Avec cette information, il est vu que les voyages non productifs sont absents, rendant les interlignes implicites et les voyages vers ou depuis les garages impossibles à déterminer sans données supplémentaires. Il serait donc attendu que les tournées commencent par le voyage numéro deux dans les données GTFS. Or, certaines tournées commencent par un numéro de

voyage plus élevé que deux. En effet, les voyages spéciaux, comme pour les institutions scolaires et pour les industries, ne sont pas indiqués. En quelque sorte, ces données GTFS présentent donc le réseau selon la perspective de l'utilisateur du réseau régulier. Aussi, les temps d'arrivée et de départ aux arrêts ne sont pas différenciés et sont précis à la minute près. Enfin, le tracé exact planifié sur le réseau routier n'est pas connu.

3.1.2 Détection des lignes-tracés et compression

Le format GTF est assez souple pour définir une ligne d'autobus de manière à ce que chaque passage à chaque arrêt de chaque voyage puisse être défini. Cela permet de définir différents tracés pour une ligne et de détailler les heures de passages aux arrêts, habituellement déterminées selon la vitesse de l'autobus et les conditions prévues de circulation. Ce standard est donc bien adapté à la modélisation [dynamique] basée sur l'horaire — ou « orientée horaire » — (*schedule-based*, tenir compte des chemins les plus courts et des horaires des véhicules).

Il est toutefois possible de résumer l'information en un ensemble de départs pour chaque ligne-tracé, puisque chaque ligne ne contient que quelques tracés possibles. Cela permet alors de compresser le fichier `stop_times`. Pour compresser ce fichier, chaque voyage est résumé en sa séquence d'arrêts. Les voyages d'une même ligne ayant la même séquence d'arrêts reçoivent alors un numéro de tracé (voir figure 3-2). À partir du fichier résultant, une liste des lignes-tracés existantes peut alors être extraite. Cette façon de résumer l'information est adaptée à une modélisation [statique] basée sur la ligne (*line-based*, sans tenir compte de l'horaire), mais peut aussi servir à résumer les voyages planifiés.

La figure 3-2 — avec des lignes et des arrêts fictifs — montre aussi que les voyages non commerciaux sont explicités. Pour ce faire, en se servant de l'identifiant (voir tableau 3-1), les voyages qui ne sont pas présents dans le GTFS sont supposés non commerciaux. Toutefois, s'il manque deux voyages consécutifs ou plus, il se peut qu'il y ait un voyage commercial spécial. Ces voyages ne sont pas inclus par conséquent. Enfin, il est supposé qu'un voyage d'injection dans le réseau commence chaque tournée et qu'un voyage d'éjection du réseau la termine.

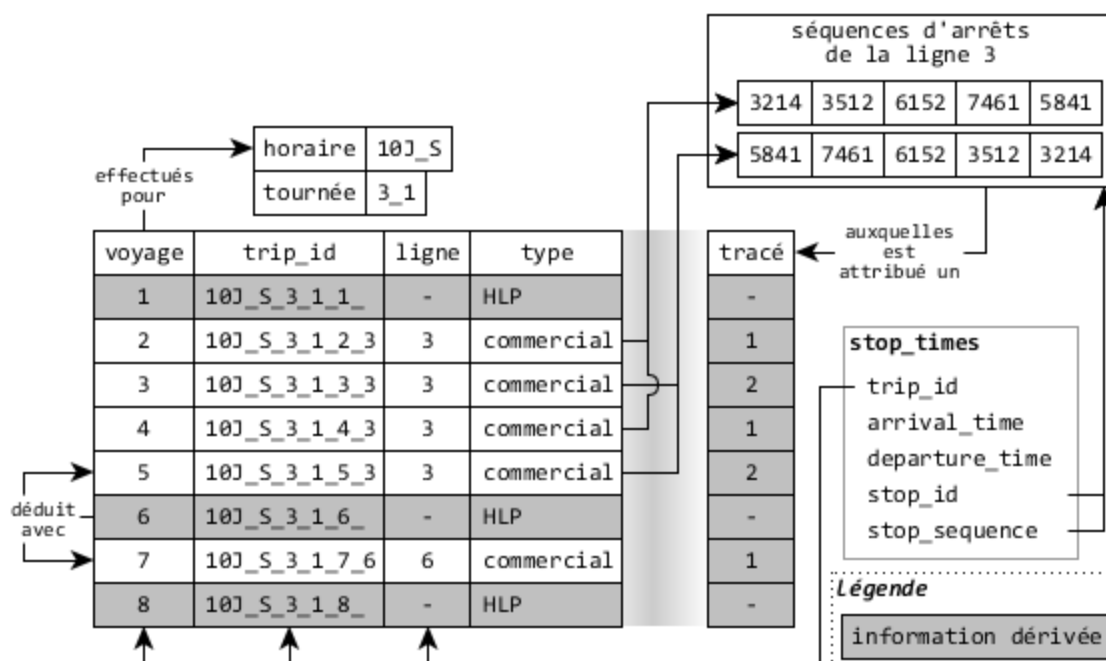


Figure 3-2. Compression des stop_times en séquences de voyages et détection des lignes-tracés : exemple avec une tournée fictive

Le tableau 3-5 montre le nombre de lignes-tracés. La très grande majorité des lignes n'ont que deux tracés, une par direction. Les autres lignes ont des tracés supplémentaires, représentant les modifications de service selon la période de la journée ou les ajustements mineurs, comme les autobus injectés en milieu du parcours principal dans une direction.

3.1.3 Aperçu du réseau planifié

Les fichiers GTFS, destinés d'abord pour de l'information voyageur, sont utilisés pour estimer et dessiner un portrait global du service d'autobus planifié. Pour ce faire, il est supposé que les heures et les lieux de départ et d'arrivée des interlignes explicités (voir section 3.1.2) correspondent respectivement à ceux du dernier arrêt du voyage précédent et à ceux du premier arrêt du voyage suivant. Les temps de battement dans une tournée sont aussi notés en supposant qu'ils surviennent entre deux voyages productifs qui sont consécutifs.

Des données sur les entrées et sorties aux garages sont utilisées pour estimer les attributs des voyages haut-le-pied depuis ou vers les garages — injection ou éjection. Ce fichier contient les heures de début et de fin des différentes tournées. Lorsque la tournée a été effectuée comme

prévue, les heures inscrites sont celles initialement planifiées. Lorsqu'une modification est apportée à la tournée, elle est inscrite en notant les nouveaux arrêts et heures de début et de fin.

Pour chaque numéro de tournée, le garage de départ qui revient le plus souvent dans le mois est pris. L'heure de début associé est notée. L'exercice est appliqué aussi pour le garage de fin de tournée. Il est donc supposé que les modifications apportées aux tournées sont rares et que la plupart des occurrences des tournées durant le mois sont effectuées avec les garages et les heures prévus. Lorsque tous les garages sont différents pour une tournée dans le mois, l'heure de départ est égal au plus tôt des départs survenus; l'heure d'arrivée, au plus tard des arrivées — ces cas ne sont pas trouvées pour les tournées dans les données GTFS.

De plus, il est supposé que les voyages HLP depuis les garages — injection dans le réseau — prennent fin au départ du premier voyage productif de la tournée et que celles vers les garages — éjection du réseau — commencent à l'arrivée au dernier arrêt du dernier voyage productif.

Une classification des lignes est utilisée en se basant sur le numéro de la ligne, calquée sur les « familles de services » de la STM (2013c), et sur ses heures d'opération — heure de pointe du matin ou de l'après-midi, midi, soir, nuit (voir tableau 3-2). Cette classification sert à décrire le service productif dans cette section.

Cette classification ne prétend pas être exhaustive, ni même la meilleure. Cette catégorisation ou hiérarchisation des lignes, qui n'est donc qu'une possibilité parmi d'autres, vise plutôt à faire ressortir, à dégager les principales tendances dans les ressources qui sont impliquées dans la mise en place du réseau en fonction des groupes ou d'une hiérarchie qui pourraient être choisis par le planificateur.

Tableau 3-2. Classification des lignes d'autobus

type			local	local, pointe	local, étendu	express	express, pointe	navette	nuit
lignes			10 à 299	10 à 299	10 à 299	400 à 499	400 à 499	700 à 799	300 à 399
périodes possibles de service (début et fin)	...	11:00	√	√	√	√	√	√	
	11:00	14:00	√		√	√		√	
	14:00	20:00	√	√	√	√	√	√	
	20:00	21:00	√	√	√	√	√	√	
	21:00	24:00	√		√	√		√	
	24:00	27:00			√	√		√	√
	27:00	...						√	√

La figure 3-3 illustre l'estimation des autobus sur le réseau planifié. Il est remarqué que le nombre d'autobus en service productif semble osciller ou vaciller dans le temps. Cela peut être dû en partie à la résolution des données à la minute, mais il faut toutefois se rappeler la définition du service productif, où les passagers sont admis à bord des autobus. En accord avec la méthodologie prise, dès le voyage productif terminé, l'autobus est en battement ou en interligne — ou en voyage HLP vers un garage. D'ailleurs, la courbe des autobus totaux sur le réseau, incluant donc les autobus en service productif ou non, est exempte de ces oscillations.

La classification utilisée pour séparer les « groupes » de lignes d'autobus montre l'ampleur des ressources ajoutées par les lignes « express » (environ 1 % à 7 % de la flotte) et par les lignes opérant en période de pointe (environ 10 %) pour le service en semaine, au moins en partie puisque les lignes « régulières » demandent aussi plus de véhicules aux heures de pointe (voir tableau 3-3 et la figure 3-3).

Tableau 3-3. Estimation des véhicules-heures pour le réseau planifié 2010 (jour de semaine)

heure	Véhicules-heures	% local étendu	% local	% express	% local pointe	% express pointe	% navette	% nuit	% battement	% interligne	% HLP garage début	% HLP garage fin
4	41.7	24.1	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.2	74.8	0.0
5	338.4	52.3	1.3	1.2	0.4	2.8	0.0	0.0	4.5	4.0	33.5	0.0
6	783.7	48.9	3.4	5.5	2.9	7.8	0.0	0.0	6.9	6.8	17.8	0.0
7	1168.6	51.0	3.6	6.6	2.7	9.3	0.3	0.0	6.6	9.6	8.6	1.7
8	1082.6	51.9	3.9	5.5	2.9	9.5	0.3	0.0	7.2	7.9	0.7	10.2
9	765.4	54.9	4.7	5.6	1.9	6.6	0.5	0.0	9.1	7.5	0.2	9.0
10	611.8	64.2	5.4	6.9	0.2	0.6	0.6	0.0	12.1	7.9	0.3	1.8
11	616.7	65.9	5.1	6.8	0.0	0.0	0.6	0.0	11.8	6.1	3.0	0.7
12	633.6	68.8	5.1	7.0	0.0	0.0	0.6	0.0	12.3	4.2	1.6	0.5
13	648.3	69.7	5.0	6.6	0.0	0.0	0.5	0.0	12.6	4.6	0.8	0.1
14	782.5	63.8	4.3	5.9	0.1	0.5	0.4	0.0	9.6	4.0	10.9	0.4
15	1131.9	52.7	3.3	4.9	1.8	4.6	0.3	0.0	7.3	4.0	18.9	2.2
16	1133.8	53.1	3.7	5.7	2.7	9.4	0.3	0.0	8.2	9.0	5.6	2.4
17	1081.6	53.0	3.9	6.4	2.6	10.0	0.3	0.0	7.9	12.1	0.4	3.3
18	894.3	53.3	3.9	5.2	2.3	7.4	0.4	0.0	7.6	9.3	0.1	10.5
19	587.3	63.1	3.1	2.8	0.6	2.1	0.6	0.0	9.3	7.6	0.2	10.6
20	437.5	72.5	1.1	0.9	0.0	0.2	0.4	0.0	11.6	10.0	0.3	2.9
21	410.6	73.3	1.1	0.5	0.0	0.0	0.4	0.0	11.4	10.2	0.4	2.5
22	362.2	74.7	0.4	0.0	0.0	0.0	0.1	0.0	11.4	7.9	0.5	5.0
23	311.3	75.7	0.0	0.0	0.0	0.0	0.0	0.0	13.5	6.6	0.4	3.9
24	268.0	72.1	0.0	0.0	0.0	0.0	0.1	0.6	13.4	6.8	1.0	5.9
25	185.0	47.9	0.0	0.0	0.0	0.0	0.0	14.4	5.6	8.0	3.3	20.8
26	68.3	2.9	0.0	0.0	0.0	0.0	0.0	60.1	11.2	8.8	3.7	13.3
27	58.1	0.0	0.0	0.0	0.0	0.0	0.0	70.7	16.9	5.6	2.9	3.8
28	46.5	0.0	0.0	0.0	0.0	0.0	0.0	76.9	14.7	1.8	0.0	6.6
29	22.0	0.0	0.0	0.0	0.0	0.0	0.0	51.9	0.8	0.0	0.0	47.4
Total	14471.1	57.8	3.5	4.9	1.4	4.7	0.3	1.1	9.1	7.5	5.7	4.1

Tableau 3-4. Estimation des voyages pour le réseau planifié 2010 (jour de semaine)

type	voyages	%
régulier étendu	13 899	54.3
régulier	1 022	4.0
régulier pointe	562	2.2
express	1 095	4.3
express pointe	821	3.2
nuit	252	1.0
navette	89	0.3
interligne	3 857	15.1
injection	1 989	7.8
éjection	1 989	7.8
total	25 575	100.0

(voyages regroupés en environ 1 989 tournées, dont 17 tournées avec données sur les garages introuvables)

Tableau 3-5. Lignes en fonction du nombre de tracés possibles (réseau 2010, jour de semaine)

tracés	lignes	exemples de lignes
1	1	21
2	121	10, 12, 13, 14, 16, 17, ...
3	41	15, 27, 29, 34, 43, 45, ...
4	24	11, 33, 52, 54, 68, 85, ...
5	4	97, 166, 194, 535
6	0	
7	1	460

(total : 192 lignes et 489 tracés)

Les voyages d'injection au réseau sont concentrés au début des périodes de pointe, alors que les voyages d'éjection du réseau sont nombreux à la fin des périodes de pointe, quoique certains de ceux-là soient plutôt répartis durant le soir et la nuit.

Il semble y avoir un nombre important d'autobus qui sont envoyés sur le réseau entre 14:00 et 16:00, alors que les voyages d'injection le matin sont plus étalés dans le temps. Le nombre élevé d'institutions scolaires qui terminent les cours à ces heures pourrait expliquer en grande partie cette injection « massive » d'autobus en après-midi.

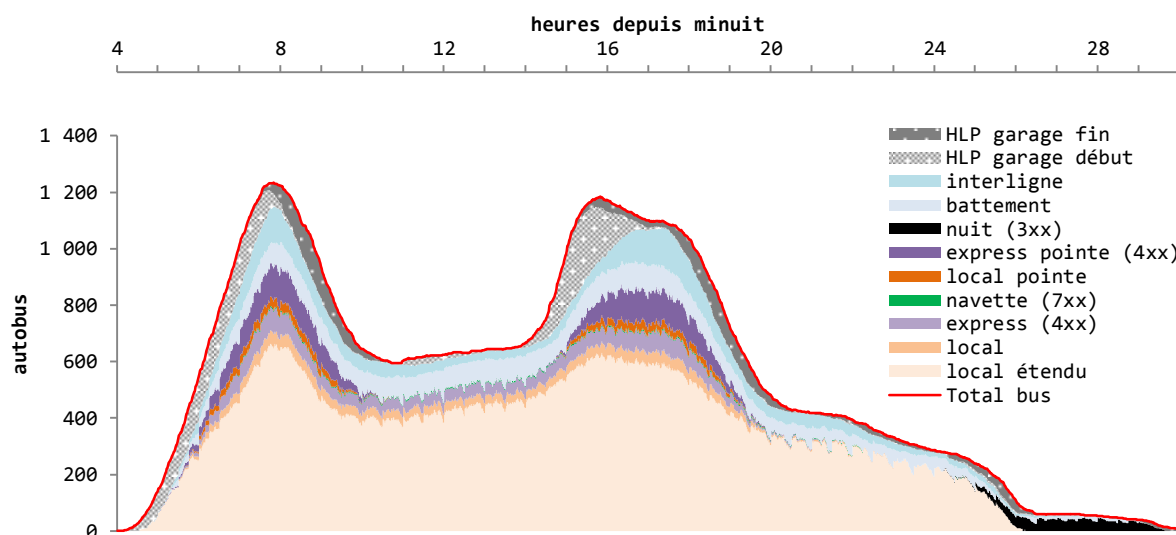


Figure 3-3. Estimation des autobus utilisés pour le réseau planifié 2010 (jour de semaine)

Il faut toutefois comprendre que les données GTFS n'affichent pas l'entièreté des tournées et des voyages prévus pour une journée. Le fichier de données sur les entrées et sorties aux garages contient plusieurs tournées « spéciales » pour les institutions scolaires ou pour des entreprises notamment, et ces voyages ne sont pas recensés dans les fichiers GTFS.

Le fichier *shapes*, indiquant la géométrie des lignes sur le réseau routier, n'étant pas disponible pour les données 2010, il a été reconstruit à partir du fichier lignes-tracés dérivé des *stop_times*. En reprenant la catégorisation utilisée précédemment, la hiérarchisation du réseau apparaît alors à la figure 3-4. La spatialisation des lignes-tracés reste à affiner, car les liens entre les arrêts sont dessinés à vol d'oiseau. Cette problématique reste très visible surtout pour les lignes express à cause de la grande distance séparant certains arrêts.

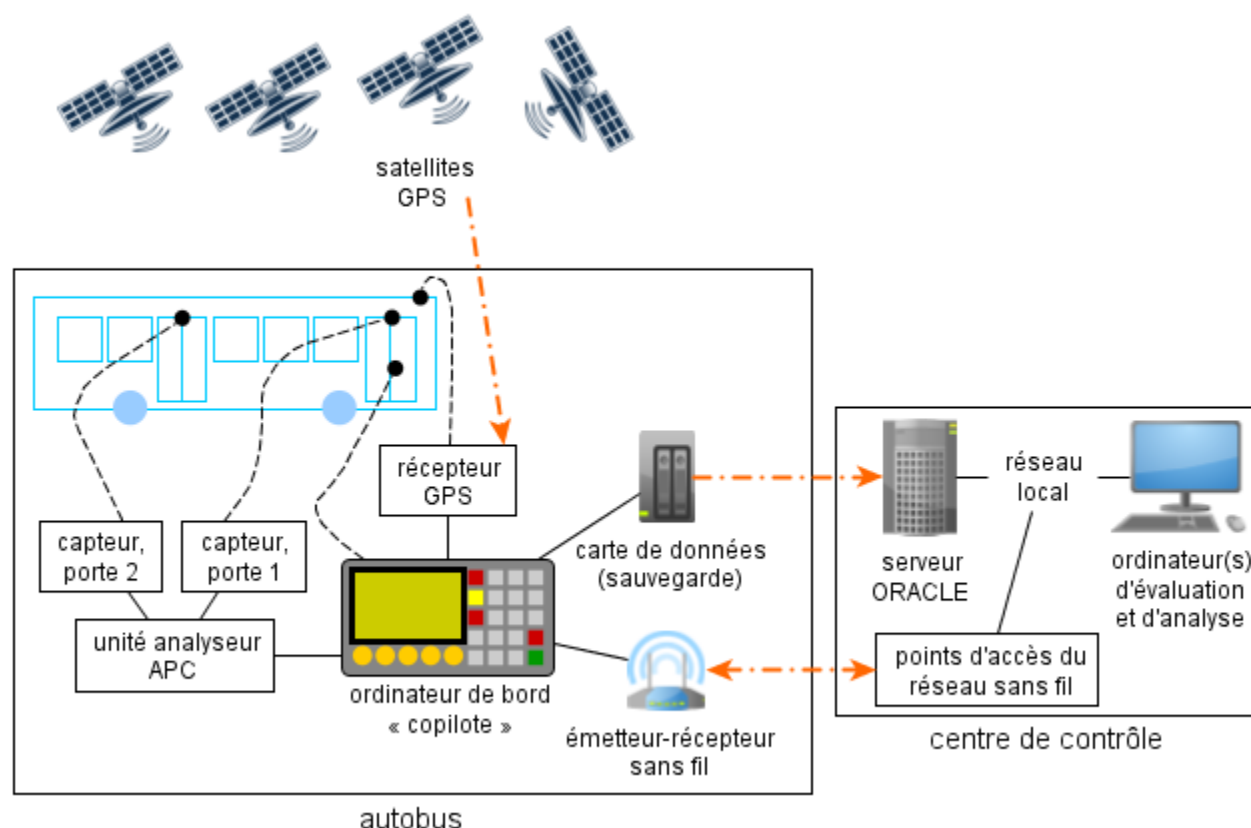


Figure 3-4. Lignes-tracés 2010, reconstruites par segments à vol d'oiseau entre les arrêts

3.2 Systèmes AVL-APC

Plusieurs autobus de la STM sont équipés de récepteurs GPS, alimentant le système AVL, et de compteurs automatiques de passagers. Ces instruments sont reliés à l'ordinateur de bord qui enregistre les données et qui peut les transmettre à un centre de contrôle. Doté d'une interface ayant notamment des cartes SIG, le centre de contrôle peut analyser les données et donner des instructions au chauffeur de l'autobus au besoin. Les données sont archivées dans les cartes de données de l'autobus et peuvent être récupérées au retour du véhicule au garage ou lors de l'entretien du véhicule.

La figure 3-5 montre de manière simplifiée le système AVL-APC installé dans plusieurs autobus de la STM et les liens entre ses diverses composantes. Ce système peut ainsi faire partie d'un SAEIV, comme dans le cas de la STM avec son projet iBus (STM, 2012d). Le système est installé par INIT, qui s'occupe aussi de faire correspondre les montées et les descentes détectées par les compteurs de passagers aux arrêts (Init Innovation in traffic systems AG [INIT], 2011).



(modifiée de INIT, 2011) (voir annexe 1 pour références des icônes)

Figure 3-5. Système AVL-APC avec un autobus régulier à deux portes

3.2.1 Données brutes

Le fichier de données combine ainsi traces GPS et comptages de montées et de descentes des usagers pour tous les véhicules équipés ou instrumentés du réseau. La distance parcourue est aussi enregistrée avec un odomètre. La position de l'autobus est déterminée par GPS toutes les quatre secondes environ (codes 2, points GPS). Après une halte, sa durée et son emplacement sont enregistrés (code 1). Si une porte a été ouverte, les données des compteurs de passagers sont notées (code 3). Combinés avec les codes 1, les codes 3 complètent ainsi l'information sur les mouvements de passagers aux haltes. Le tableau 3-6 montre les attributs enregistrés pour chacun des événements, alors que la figure 3-6 affiche sur une carte ces différents événements avec un extrait d'une trace GPS d'un véhicule.

Tableau 3-6. Attributs possibles pour chaque entrée du fichier de données du système AVL-APC

attribut	point GPS	haltes	comptage	valeurs possibles	(notes)
véhicule	√	√	√	entier positif	
code	2	1	3	1, 2 ou 3	type d'enregistrement
temps	√		√	entier positif	secondes depuis minuit
temps début		√		entier positif	
temps fin		√		entier positif	
longitude	√	√		nombre réel	système géodésique WGS84
latitude	√	√		nombre réel	
distance	√	√		entier positif	selon l'odomètre, en mètres
état des portes		√		0 ou 1	0 : portes fermées 1 : au moins une porte ouverte
numéro de porte			√	entier positif	
montées			√	entier positif	
descentes			√	entier positif	

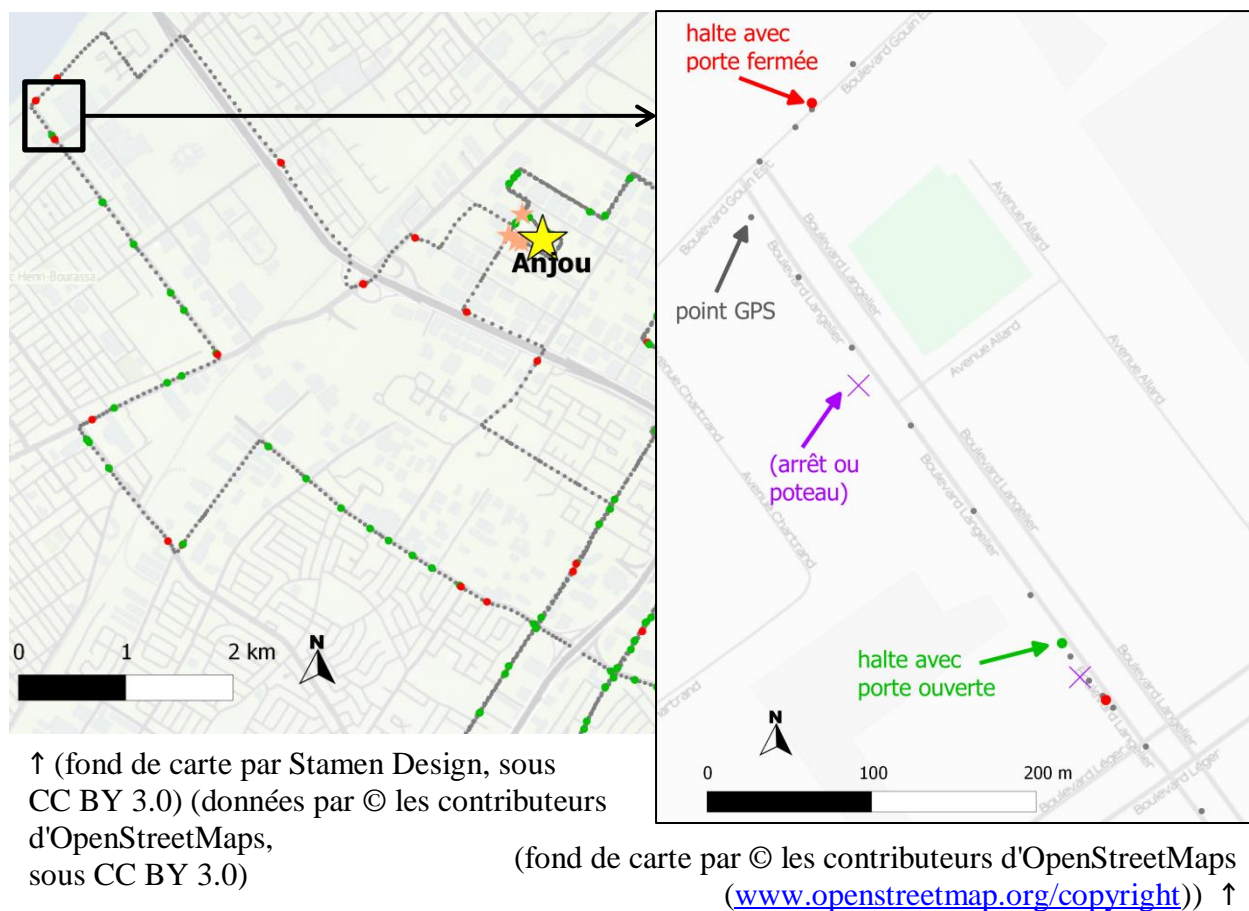


Figure 3-6. Aperçu d'une partie de la trace GPS du véhicule 21211

Des données pour 275 autobus sont disponibles pour le mercredi 17 mars 2010, ce qui représente environ 16.5 % de la flotte en opération durant mars 2010 — en se basant sur les 1 669 autobus présents dans le fichier d'entrées et de sorties aux garages (déjà mentionné à la section 3.1.3).

L'entreprise qui prend en charge la collecte des données des systèmes AVL-APC — INIT, inc. — s'occupe aussi de leur traitement. Les heures de passage — arrivée et départ — et les mouvements des usagers ont donc été associés à chaque arrêt, produisant ainsi les « arrêts réels » et les « voyages réels » — dans le sens qu'ils s'approchent des passages effectivement réalisés. Les données de comptage sont aussi balancées. Les algorithmes liés à ces opérations ne sont pas accessibles, mais leurs résultats permettent d'en dégager les grandes lignes.

3.2.2 Détection des arrêts

En général, les temps de passage aux arrêts sont détectés à l'aide des points GPS (codes 2) lorsque l'autobus ne s'y arrête pas et avec les haltes (codes 1) dans le cas contraire. La distance (depuis le terminus) de l'odomètre est aussi enregistrée. Toutefois, les temps d'arrivée et de départ au terminus d'arrivée d'un voyage sont les mêmes que pour le terminus de départ du voyage suivant. Donc, le temps d'arrivée au premier arrêt et le temps de départ du dernier arrêt du voyage ne sont pas considérés. De plus, les temps de passage ne sont pas enregistrés aux arrêts où le signal GPS est interrompu. D'ailleurs, le système semble aussi abandonner la détection d'un voyage complet si le premier arrêt n'est pas détecté.

Ainsi, 3 812 voyages réels, effectués par 233 véhicules, sont recensés. En utilisant l'identifiant de voyage (voir tableau 3-1), environ 3 567 voyages semblent correspondre aux voyages GTFS, appartenant à 297 tournées d'autobus et 227 véhicules. Le tableau 3-7 montre la répartition de ces voyages selon leur type et le taux de détection des voyages pour les tournées ayant au moins un voyage réel. Les 245 voyages non recensés dans les données GTFS comprennent notamment des voyages pour les écoliers et des industries et des voyages spéciaux ajoutés et non prévus pour ajuster l'offre service.

Tableau 3-7. Estimation des voyages réels associés aux voyages GTFS

type	total	détectés	%	tournées détectées	
				voyages totaux	%
productives	17 740	2 601	14.7	3 005	86.6
interlignes	3 857	478	12.4	673	71.0
injection	1 989	252	12.7	297	84.8
éjection	1 989	236	11.9	297	79.5
total	25 575	3 567	13.9	4272	83.5

La figure 3-7 reprend l'estimation des autobus et des véhicules-heures en fonction du temps en ne conservant que les voyages GTFS qui ont été associés à un voyage réel. Les classes de lignes sont les mêmes que celles utilisées à la section 3.1.3. Les autobus en battement sont pris en compte seulement lorsque les voyages précédents et suivants sont détectés. Il est remarqué que les classes de lignes ne sont pas représentées dans les mêmes proportions que dans les voyages GTFS. Tous ces voyages représentent environ 2 110.8 véhicules-heures du réseau planifié (voir annexe 3 pour les détails).

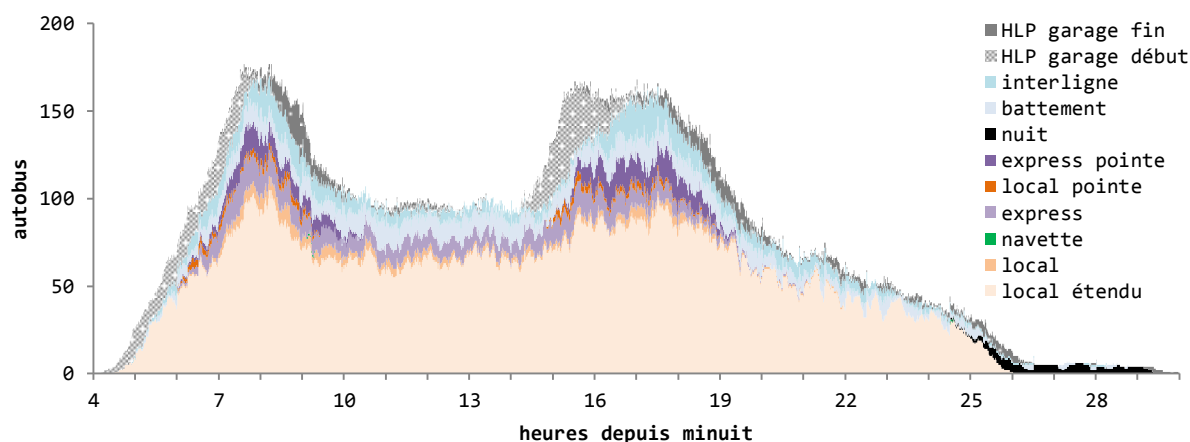


Figure 3-7. Estimation des autobus en fonction du temps sur le réseau planifié, en conservant seulement les voyages GTFS associés à un voyage réel

3.2.3 Balancement des mouvements de passagers

La figure 3-8 affiche la distribution des véhicules selon la différence entre les montées et les descentes mesurées par les compteurs de passagers avant leur balancement.

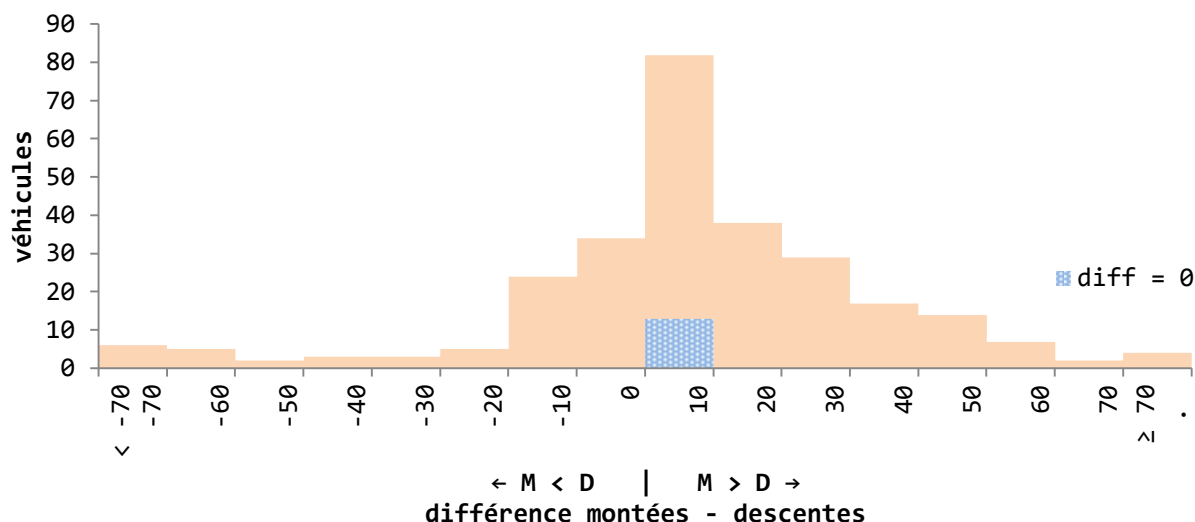


Figure 3-8. Distribution des véhicules selon la différence montées-descentes

Les véhicules n'ayant pas de différence entre ces deux types de mouvements sont rares. La majorité des véhicules ont entre un déficit d'une vingtaine de montées et un surplus d'une quarantaine de montées. En effet, le comptage des passagers comporte fréquemment un débalancement entre les montants et les descendants. Bien que cela n'affecte pas nécessairement la planification et les analyses opérationnelles de manière importante lorsque ces différences sont mineures, la cohérence de ces données reste importante pour conserver leur crédibilité (Kikuchi, Mangalpally, & Gupta, 2006).

Le tableau 3-8 montre le résultat du balancement des mouvements de passagers selon les lignes, en montrant celles qui présentent des voyages partiellement balancés. Globalement, le balancement des mouvements laisse un surplus de 1 195 montées, soit environ 0.8 %, ce qui reste très faible. La plupart des lignes présentent un débalancement nul ou faible. Les lignes 69, 121 et 467 présentent un débalancement très élevé en valeur absolue mais pas en valeur relative, car ce sont des lignes très achalandées (STM, 2010a, 2011a). La non-détection du terminus d'arrivée de la ligne 73 (station Du-Collège) semble expliquer son débalancement relatif élevé entre les mouvements de passagers.

Tableau 3-8. Différence entre les montées et les descentes « balancées »

ligne	montées	descentes	différence	%
11	1 056	1 056	0	0.0
15	106	110	-4	-3.8
36	678	682	-4	-0.6
45	2 954	2 955	-1	0.0
69	8 988	8 571	417	4.6
73	110	66	44	40.0
115	122	121	1	0.8
121	8 930	8 656	274	3.1
132	2 393	2 403	-10	-0.4
144	2 248	2 262	-14	-0.6
168	581	577	4	0.7
170	1 040	1 032	8	0.8
179	773	773	0	0.0
467	6 129	5 649	480	7.8
autres	113 732	113 732	0	0.0
total	149 840	148 645	1 195	0.8

3.3 Système AFC

L'utilisateur doit acquitter son droit de passage à bord des autobus, ce qui est vérifié par le système AFC — il y a ainsi une validation du passage. Dans les autobus, ce système est confiné à la boîte de perception, situé à côté du chauffeur, près de la porte avant. La carte à puce OPUS a été introduite au sein de la STM à l'automne 2008 et généralisée en 2009 (STM, 2012a). Perçu d'abord comme un moyen de mieux contrôler les recettes par la réduction de la fraude (STM, 2008), le système représente une grande opportunité pour la planification puisqu'il est possible, théoriquement, de suivre à la trace les déplacements de chaque usager du réseau pour la durée de vie utile de la carte.

3.3.1 Acquittement du droit de passage

Les usagers peuvent payer comptant ou utiliser un titre de transport sur un billet à bande magnétique ou sur une carte à puce OPUS. Lorsque l'utilisateur paye comptant, il obtient en retour un billet de correspondance pour conserver une preuve de paiement selon le règlement R-105 de la STM (2013a). L'émission de ce billet est enregistrée, le sous-système enregistrant la monnaie étant séparé du système AFC. Toutefois, l'acquittement du droit de passage n'est pas requis pour plusieurs personnes, comme les enfants de moins de six ans « accompagné d'une personne en assumant la surveillance » et les policiers et les pompiers en uniforme (STM, 2013a).

Sur le réseau de la STM, il y a possibilité de correspondre sur une ou plusieurs autres lignes dans les 120 minutes après l'entrée dans le réseau. L'utilisateur ne peut pas retourner dans le métro avec le même passage si une validation a été effectuée dans le métro auparavant. La validation du titre de transport est nécessaire lors d'une correspondance, sauf lorsque l'utilisateur passe d'une ligne de métro à une autre. Il n'y a toutefois pas de validation à effectuer à la descente du véhicule.

3.3.2 Traitements

Les données brutes de validations de passages du 17 mars 2010 contiennent les séries de validations de passages enregistrées aux boîtes de perception dans les autobus et aux tourniquets ou aux portillons dans le métro. Les boîtes de perception enregistrent aussi les erreurs, comme celles causées par une carte à puce illisible ou par un titre de transport non accepté. Toutes ces données sont récupérées au retour de l'autobus au garage ou lors de son entretien. D'anciens travaux, comme celui de Doucet-Koussaya (2012) à propos du réseau de métro, ont été réutilisés pour avoir une idée de la signification des différents attributs des validations de passages et pour traiter préliminairement les données.

Ainsi, les enregistrements doubles, les refus, les activités d'essai — types d'activités « BF » ou « BG », identifiant le début d'une série de validations — et les validations tests ont été retirés de la base de données. Il reste alors environ 1.7 million de validations (tableau 3-9) :

Tableau 3-9. Validations de passages dans la base de données

type	10 ³ entrées	%
conservées	1 692.6	48.0
doublons	1 762.3	50.0
refus	45.6	1.3
activités BF, BG	21.8	0.6
tests	2.3	0.1

3.3.3 Validations

Les validations à l'entrée du réseau et à la correspondance sont différenciées dans la base de données. La correspondance est ici considérée selon l'opérateur comme un changement de ligne dans les 120 minutes suivant la première validation ce qui diffère de la définition de correspondance pour le planificateur — un changement de véhicule lors d'un déplacement.

La figure 3-9 et le tableau 3-10 montrent les validations enregistrées dans les autobus et dans les stations de métro en discriminant les validations à l’entrée du réseau de la STM et celles considérées aux correspondances selon le système. La figure utilise une résolution temporelle à la minute, alors que le tableau indique le total des validations dans la journée.

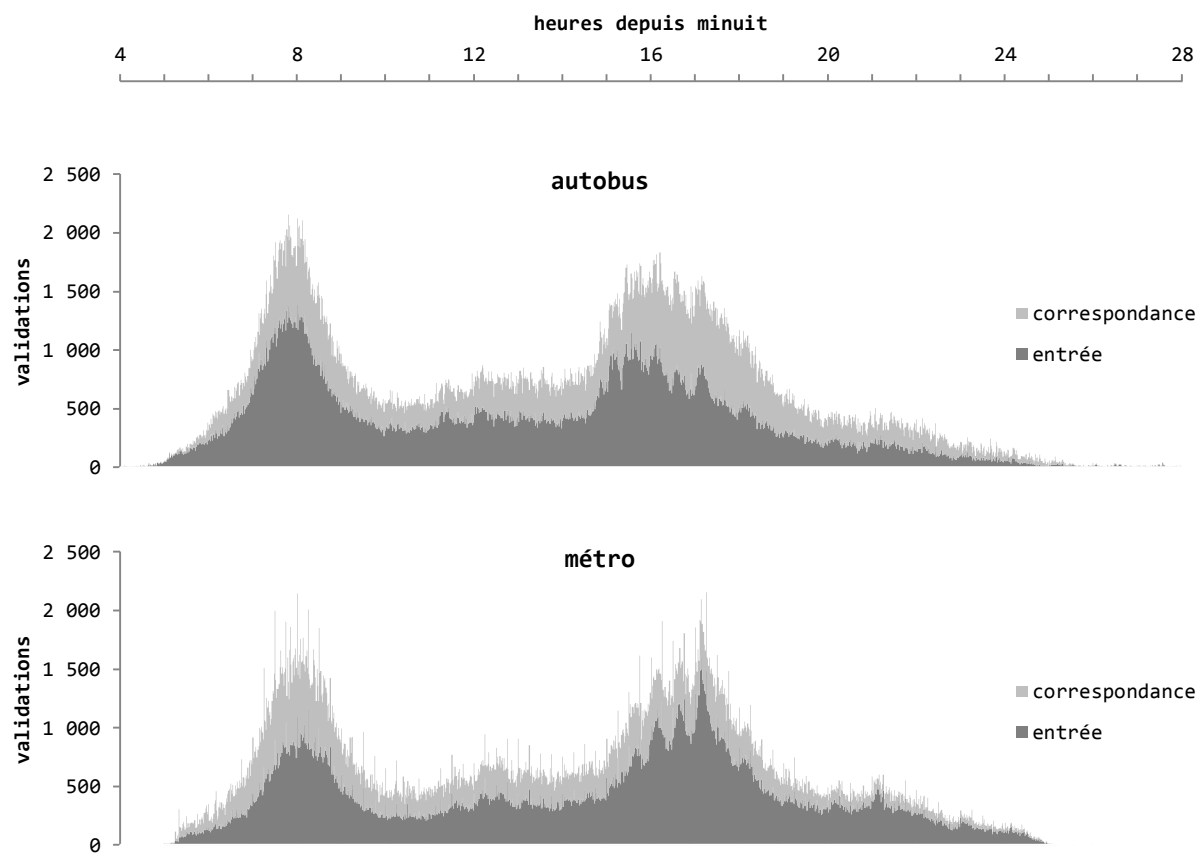


Figure 3-9. Validations de titres en fonction du mode et du type d’activité dans le temps

Tableau 3-10. Validations de passages par mode et par type d’activité

10 ³ validations		mode		total
		autobus	métro	
type	entrée	518.5	505.8	1024.3
	correspondance	376.2	292.1	668.3
	total	894.7	797.9	1692.6

%		mode		total
		autobus	métro	
type	entrée	30.6	29.9	60.5
	correspondance	22.2	17.3	39.5
	total	52.9	47.1	100.0

Les importantes variations, ou les pics, dans le graphique des validations dans le métro à la figure 3-9 sont probablement dues, notamment, au fait qu'il y a plusieurs tourniquets dans une station : les validations de passages s'accumulent ainsi plus rapidement que dans un autobus. Les personnes qui arrivent en masse des autobus et qui correspondent au métro peuvent aussi expliquer ces pics.

Il faut toutefois souligner que les correspondances dans la base de données incluent les usagers qui proviennent d'autres sociétés de transport. Par exemple, un usager ayant pris un autobus à Laval, sur le réseau de la Société de transport de Laval (STL), et qui correspond à la station de métro Cartier n'aura qu'une seule validation dans la base de données, soit la correspondance au métro. En quelque sorte, la base de données couvre l'ensemble des validations de passages sur le réseau de la STM et tient compte des correspondances entre les réseaux des autorités de transport de la région.

Chaque validation de passage dans la base de données est liée à la boîte de perception qui l'a enregistrée et au voyage dans lequel l'usager s'est trouvé. Normalement, un chauffeur d'autobus doit sélectionner et valider le voyage qu'il s'apprête à effectuer en spécifiant entre autres la ligne et la direction. Or, il se peut que le chauffeur n'accomplisse pas cette opération, ce qui compromet la cohérence des données sur le voyage. Pour contourner le problème, avec les données AVL-APC, la section 4.3 propose une méthode d'appariement des données de validations aux arrêts.

CHAPITRE 4 EXPLORATION ET TRAITEMENT DES DONNÉES

Ce chapitre expose le développement d'outils visant une intégration des données GTFS et celles des systèmes AVL-APC et AFC. D'abord, les corridors qui serviront d'exemple d'application sont choisis. La visualisation de certains indicateurs par boîtes à moustache est discutée ensuite. Après, les validations de passages sont appariées aux arrêts en vue de leur confrontation aux données de comptage. Puis, la ponctualité des autobus en début et en fin de voyage est vue. Enfin, un diagramme espace-temps interactif et un tableau de bord résument les activités sur les corridors d'autobus.

4.1 Choix des corridors

La notion de corridor peut être définie comme une bande plus ou moins large de territoire suivant un flot général et directionnel d'usagers ou reliant des générateurs de déplacements importants (Transportation Research Board & Kittelson and Associates Incorporated, 2003). Smith (1999) considère un corridor comme un regroupement, linéaire spatialement, de déplacements — réalisés ou potentiels — dans une région desservie par des services de transport (routes, lignes de transport en commun, pistes cyclables, etc.) alimentant des liens de plus grande importance (autoroute, services d'autobus express, etc.).

Toutefois, la notion de corridor reste encore vague. D'ailleurs, Douglass et Dryden (2012) constatent que le terme « corridor » a plusieurs sens selon différentes personnes. D'autres questions peuvent être soulevées, comme si une ligne ne traversant que partiellement le corridor peut être considéré comme une ligne du corridor ou non. En somme, un corridor désigne une région desservie par une ou plusieurs lignes importantes de transport. Ce mémoire n'insiste pas sur le territoire autour des lignes d'autobus, mais se concentre plutôt sur les lignes elles-mêmes.

Les corridors pris comme exemples pour les prochains chapitres ont été choisis en fonction de l'importance des lignes qui y sont présentes mais aussi en fonction du nombre de voyages « instrumentés » — dont les données sont disponibles. Le tableau suivant montre les corridors ayant le plus de voyages détectés :

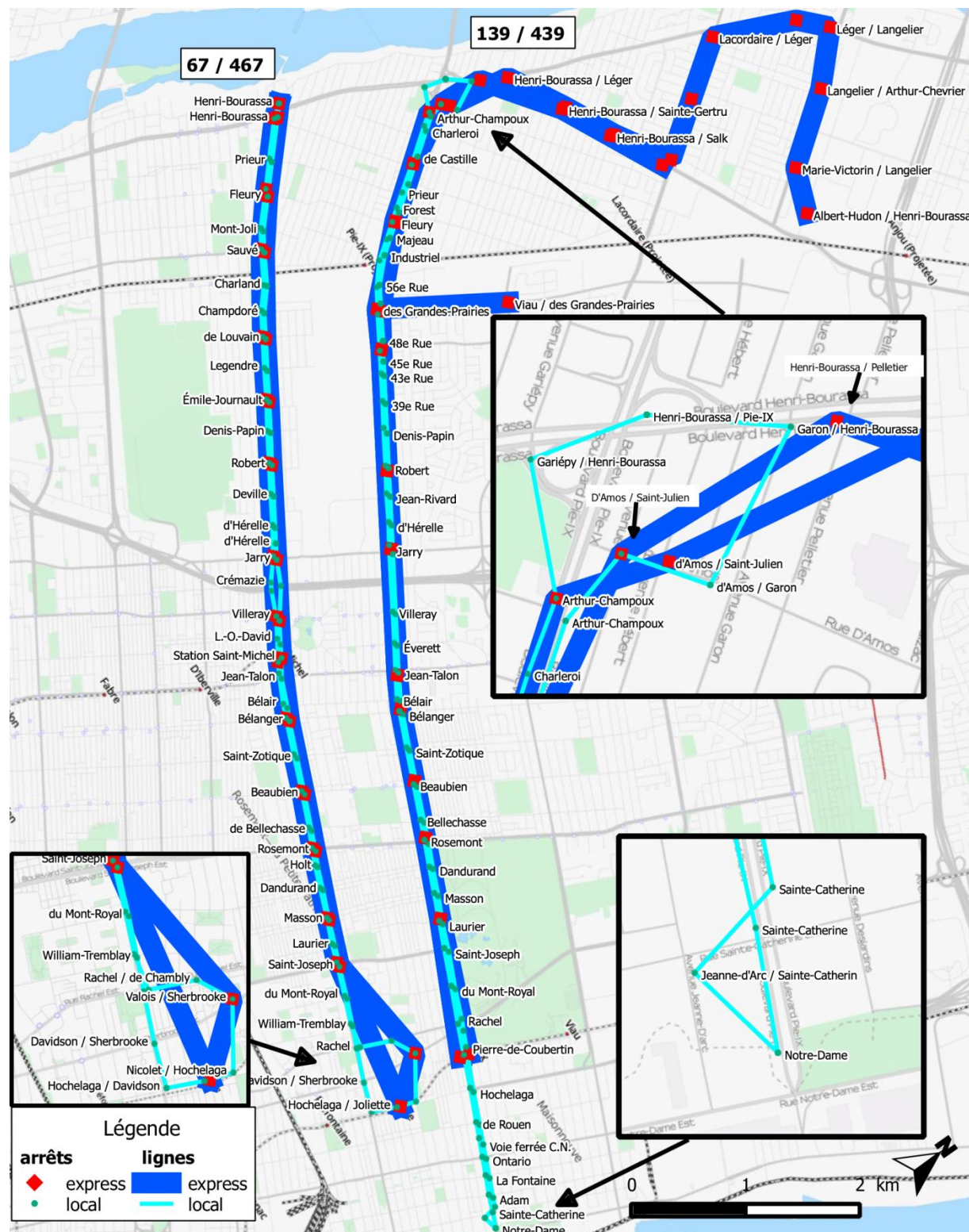
Tableau 4-1. Corridors ayant le plus de voyages détectés

lignes	(numéros depuis 2012)	nom	voyages détectés	%
139 / 505	139 / 439	Pie-IX	173	35.6
69 / 159	69 / 469	Gouin/Henri-Bourassa	153	37.5
67 / 467	67 / 467	Saint-Michel	115	22.9
121	121	Sauvé / Côte-Vertu	80	23.9
97	97	Mont-Royal	59	35.1

Les corridors 139/505 et 67/467 sont parmi les plus achalandés du réseau de la STM (2011a) et sont aussi parmi les plus instrumentés. Ainsi, les corridors Saint-Michel et Pie-IX, respectivement pour les lignes 67/467 et 139/439, sont donc pris pour ce mémoire. La figure 4-1 montre les corridors et l'emplacement des arrêts de leurs lignes sur une carte. Le trajet des lignes y est tracé à vol d'oiseau entre les arrêts.

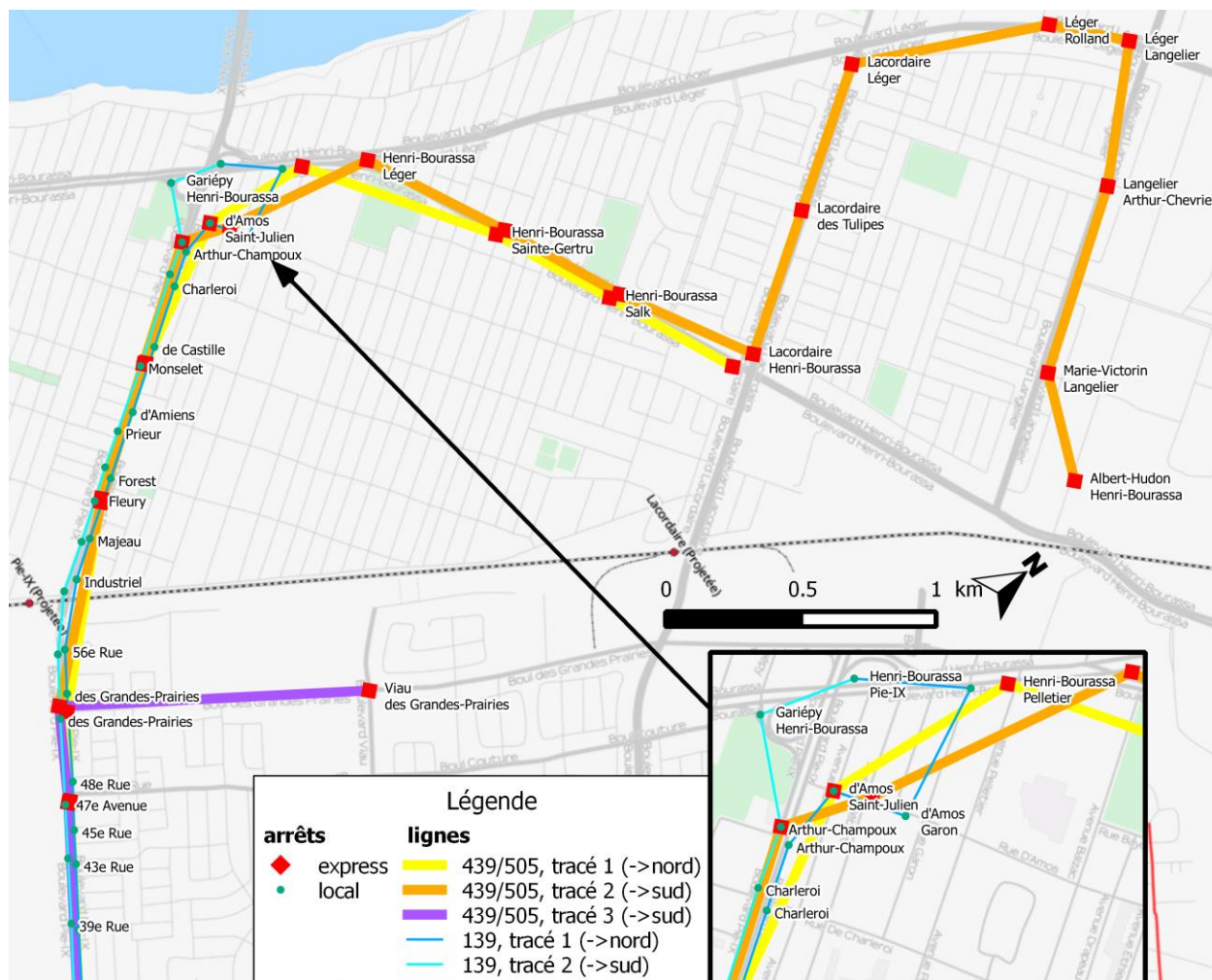
Les lignes 67 et 467 empruntent le boulevard Saint-Michel et relient les stations Saint-Michel et Joliette (intersection Hochelaga/Joliette) à l'intersection Saint-Michel/Henri-Bourassa. Les lignes s'arrêtent notamment à la station Saint-Michel. Les deux lignes partagent essentiellement le même trajet. Pour se rendre à la station Joliette, les autobus de ces lignes font une boucle sur les rues Rachel, Valois, Sherbrooke, Nicolet, Hochelaga et Davidson.

Le service de la ligne 139 longe le boulevard Pie-IX, connecte le boulevard Henri-Bourassa à la station Pie-IX et se termine à l'intersection avec la rue Notre-Dame. L'express 439, anciennement la 505, ne se rend qu'à la station Pie-IX au sud, mais son service au nord permet de desservir l'est de l'arrondissement Montréal-Nord. Quelques départs de cette ligne commencent à Viau/des Grandes-Prairies, plus au sud. La figure 4-2 zoome sur la partie nord des tracés de la ligne 439.



(fond de carte par © les contributeurs d'OpenStreetMap (www.openstreetmap.org/copyright))

Figure 4-1. Corridors des lignes 67/467 et 139/439



(fond de carte par © les contributeurs d'OpenStreetMap (www.openstreetmap.org/copyright))

Figure 4-2. Corridor Pie-IX (tracés des lignes 439/505)

4.2 Représentation de certains indicateurs

Certains indicateurs sont représentés par leur distribution avec des boîtes à moustaches (BAM), comme la ponctualité à un arrêt ou la durée d'une halte à un arrêt. Au lieu de réduire un phénomène à une mesure statistique, comme une moyenne ou un écart-type, représenter la distribution permet de mieux refléter la complexité d'un phénomène. Pour ce faire, les boîtes à moustaches aident à dégager rapidement la dispersion et l'asymétrie des données ainsi que les valeurs aberrantes ("Box plot," s.d.). La figure 4-3 montre les centiles utilisés pour les BAM dans cette section.

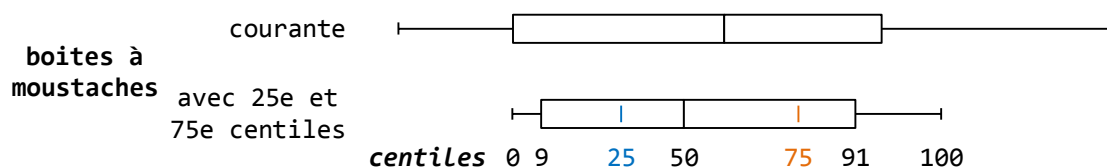


Figure 4-3. Exemples de boîtes à moustaches et centiles utilisés

Ces centiles sont utilisés en référence et en comparaison à la distribution normale, même si certains pourraient préférer les premier et neuvième déciles au lieu des 9^e et 91^e centiles pour éviter l'hypothèse d'une telle distribution sur les résultats ("Seven-number summary," s.d.). Par contre, puisque ces centiles sont proches, la représentation retenue reste tout de même valable.

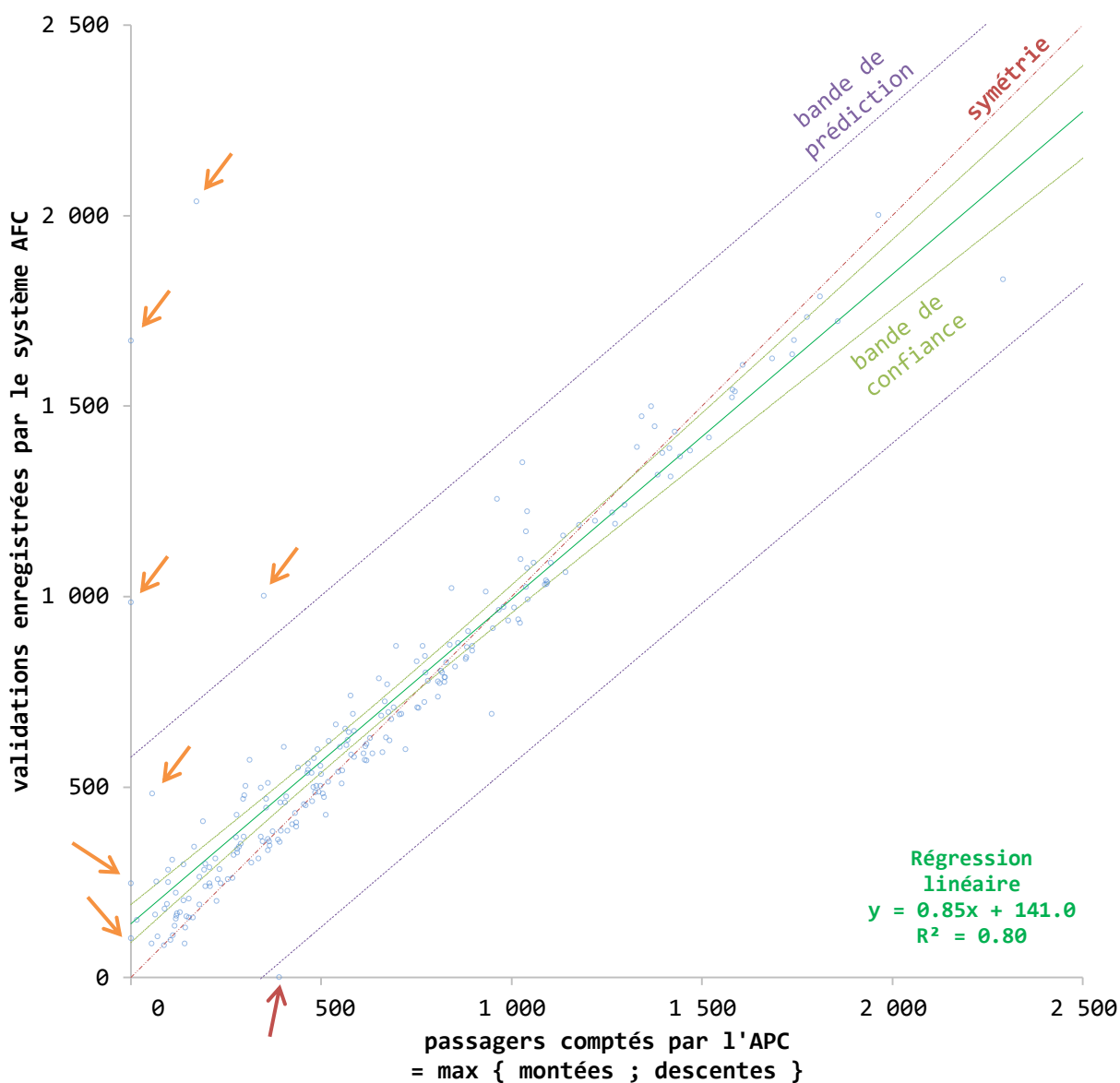
4.3 Confrontation AFC-APC

Il est possible d'estimer le nombre d'utilisateurs à partir des données AFC et APC. Toutefois, il faut se rappeler que ces systèmes mesurent fondamentalement deux actions différentes. Les compteurs de passagers mesurent les mouvements — montées et descentes — aux portes de l'autobus, alors que les boîtes de perception recensent les validations de passages. Il devient alors possible que des différences apparaissent entre les estimations basées sur ces systèmes. Par exemple, un usager qui monte à bord d'un autobus sans valider son titre de transport est compté par l'APC mais pas par l'AFC.

4.3.1 Modèles de régression et appariement aux arrêts

La figure 4-4 résume le modèle de régression linéaire confrontant les données brutes des compteurs automatiques de passagers et celles des boîtes de perception pour chaque véhicule. Certains véhicules semblent avoir des défaillances, à cause des compteurs ou de la boîte de perception (voir flèches sur la figure). De plus, la bande de confiance du modèle de régression — estimant en quelque sorte la précision avec laquelle la droite est déterminée en tenant compte de l'intervalle de confiance sur les coefficients — exclut l'axe de symétrie. La bande de prédiction du modèle — estimant les limites dans lesquelles une nouvelle observation tomberait en supposant une même population statistique — inclut la grande majorité des observations et aussi

l'axe de symétrie dans le domaine illustré dans le graphique. Par contre, la grande largeur de cette bande indique que les erreurs affectent beaucoup la précision du modèle.



modèle de régression		$\hat{y}_i = \beta_1 x_i + \beta_0$
pente de la droite	β_1	0.85 ± 0.06
constante	β_0	141.0 ± 49.9
coefficient de détermination	R^2	0.80
taille de l'échantillon (véhicules)	n	227
coefficient de corrélation estimé	$r = \hat{\rho}$	0.90
seuil de confiance utilisé	$1 - \alpha$	0.95

Figure 4-4. Modèle de régression linéaire confrontant les données brutes APC-AFC par véhicule

Il faut toutefois se rappeler que ce modèle confronte les données brutes pour chaque véhicule, ce qui peut biaiser le modèle étant donné, notamment, que les comptages de passagers ont été balancés et que certains voyages non détectés par les données AVL-APC sont inclus. Pour améliorer le modèle, un appariement des données AFC aux arrêts a été effectué et les voyages sans données AVL-APC ou AFC sont exclus. Ainsi, une meilleure comparaison des estimations des passagers entre les deux systèmes pourra être effectuée.

La méthodologie d'appariement de ces données aux arrêts est en partie inspirée de celle d'Eom, Choi et Lee (2012), qui ont fait correspondre des validations de CAP — effectuées à l'entrée et à la sortie des autobus — aux arrêts d'une ligne d'autobus de Séoul en Corée du Sud. L'idée générale reste de supposer qu'un usager valide son passage entre l'arrivée de l'autobus à l'arrêt et l'arrivée au prochain arrêt. En effet, il est possible d'imaginer que le chauffeur poursuive sa route alors que des usagers n'ont pas encore validé leur titre. Le temps et le numéro de boîte de perception, normalement associé à un autobus, disponibles pour chaque validation sont alors utilisés.

Pour raffiner l'hypothèse, le « prochain arrêt » est soit le prochain arrêt avec ouverture des portes ou avec montées d'usagers selon le compteur de la porte avant, soit le terminus d'arrivée. Afin de simplifier l'appellation, ces arrêts sont des « jalons ». Ils servent de repères pour délimiter les différents temps pour l'appariement des validations de titres.

De plus, il est aussi supposé que personne ne monte au dernier arrêt d'un voyage. L'appariement au premier ou au dernier arrêt des voyages ne semble donc pas poser problème malgré la détection des temps d'arrivée et de départ décrite à la section 3.2.2. En effet, puisque ce sont les temps d'arrivée qui sont pris en compte, les validations avant le départ au premier arrêt y sont associées et non au dernier arrêt du voyage précédent. Toutefois, il se peut que, sur une ligne-tracé particulière, les chauffeurs ne s'arrêtent pas au terminus prévu, l'arrêt précédent jouant alors le rôle de terminus, surtout si ces deux arrêts sont très proches. Donc, s'il n'y a systématiquement aucune détection de halte au dernier arrêt prévu d'une ligne-tracé, c'est l'arrêt précédent qui est considéré comme terminus, à condition qu'il soit à moins de 350 mètres du terminus prévu. Ce seuil a été choisi pour exclure les terminus très éloignés de l'arrêt précédent, comme pour les services effectuant la navette entre un quartier et une station de métro éloigné —

comme la ligne 73 en direction de la station Du Collège. Dans ces cas, il est beaucoup plus probable qu'il y ait un problème de détection de l'arrêt.

Enfin, s'il y a un « trou » de détection d'arrêt par données AVL entre deux arrêts jalons, il est supposé que le temps de fin d'appariement au premier arrêt est égal au temps de départ à l'arrêt précédent immédiatement le trou, qu'il soit jalon ou non. Les conditions d'appariement sont résumées à l'aide d'une ligne de temps à la figure 4-5.

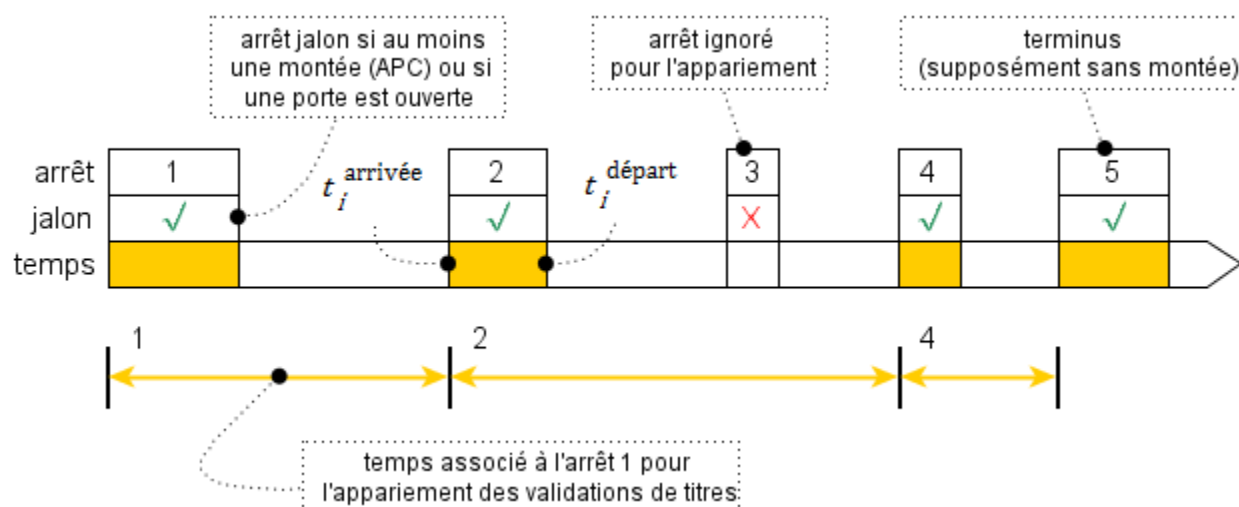


Figure 4-5. Arrêts jalons pour l'appariement des données AFC aux arrêts d'un voyage productif

Les validations de titres sont ensuite comparées aux montées selon les données balancées des compteurs de passagers. Les voyages non détectés avec les données AVL sont écartés, de même que les véhicules sans aucune donnée AFC. Pour exclure les déficiences les plus flagrantes des instruments, les voyages sans validation de titre ou sans montées selon les compteurs sont aussi exclus — 23 voyages ont ce problème, touchant 15 véhicules différents, et sont presque tous des cas isolés dans la tournée du véhicule.

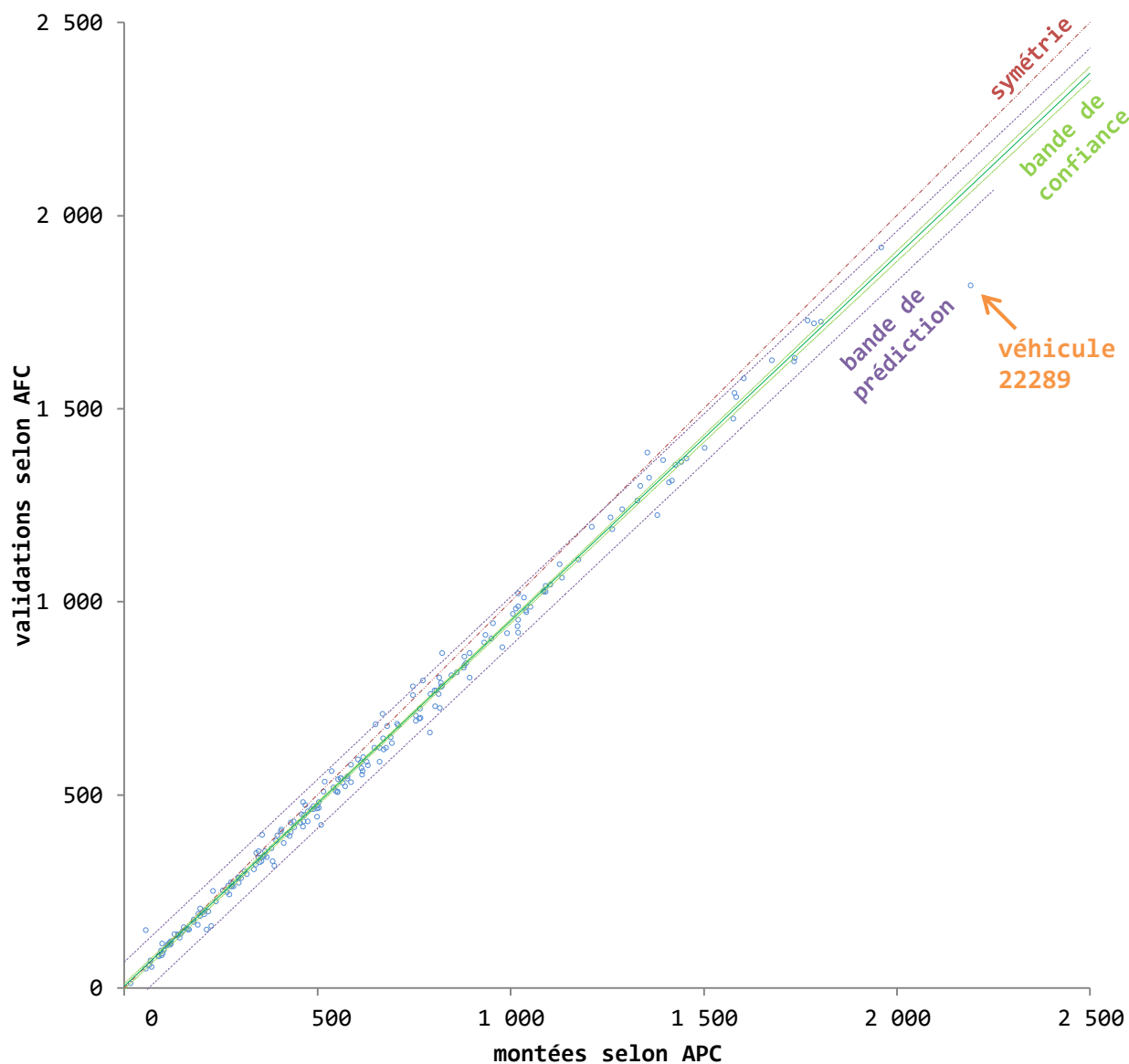
En tout, 143 488 montées et 136 642 validations ont été retenues, soit une différence de 6 535 montées ou validations — environ 4.8 % par rapport aux comptages. Les compteurs tendraient donc à estimer plus de passagers que les boîtes de perception. Les causes les plus probables incluent entre autres un dysfonctionnement des instruments — les données retirées, comme vu au paragraphe précédent, ne concernent que les instruments qui ne fonctionnent pas du

tout —, des usagers n’ayant pas correctement validé leur titre de transport ou des usagers sortant brièvement de l’autobus pour laisser descendre d’autres usagers.

Il faut toutefois souligner que les usagers ayant payé par argent comptant ne sont pas directement inclus dans les données AFC. Ils le sont cependant indirectement par le billet de correspondance émis après avoir mis la monnaie dans la boîte de perception. Néanmoins, même si le chauffeur doit émettre le billet de correspondance pour que l’usager ait au moins une preuve de paiement selon le règlement R-105 de la STM (2013a), il n’est pas garanti que cette opération soit exécutée automatiquement et systématiquement en tout temps. Un dysfonctionnement n’est donc pas à exclure de ce côté. De plus, plusieurs personnes n’ont pas à valider leur titre de transport, comme les enfants de moins de six ans « accompagné d’une personne en assumant la surveillance » ou les policiers et les pompiers en uniforme (STM, 2013a).

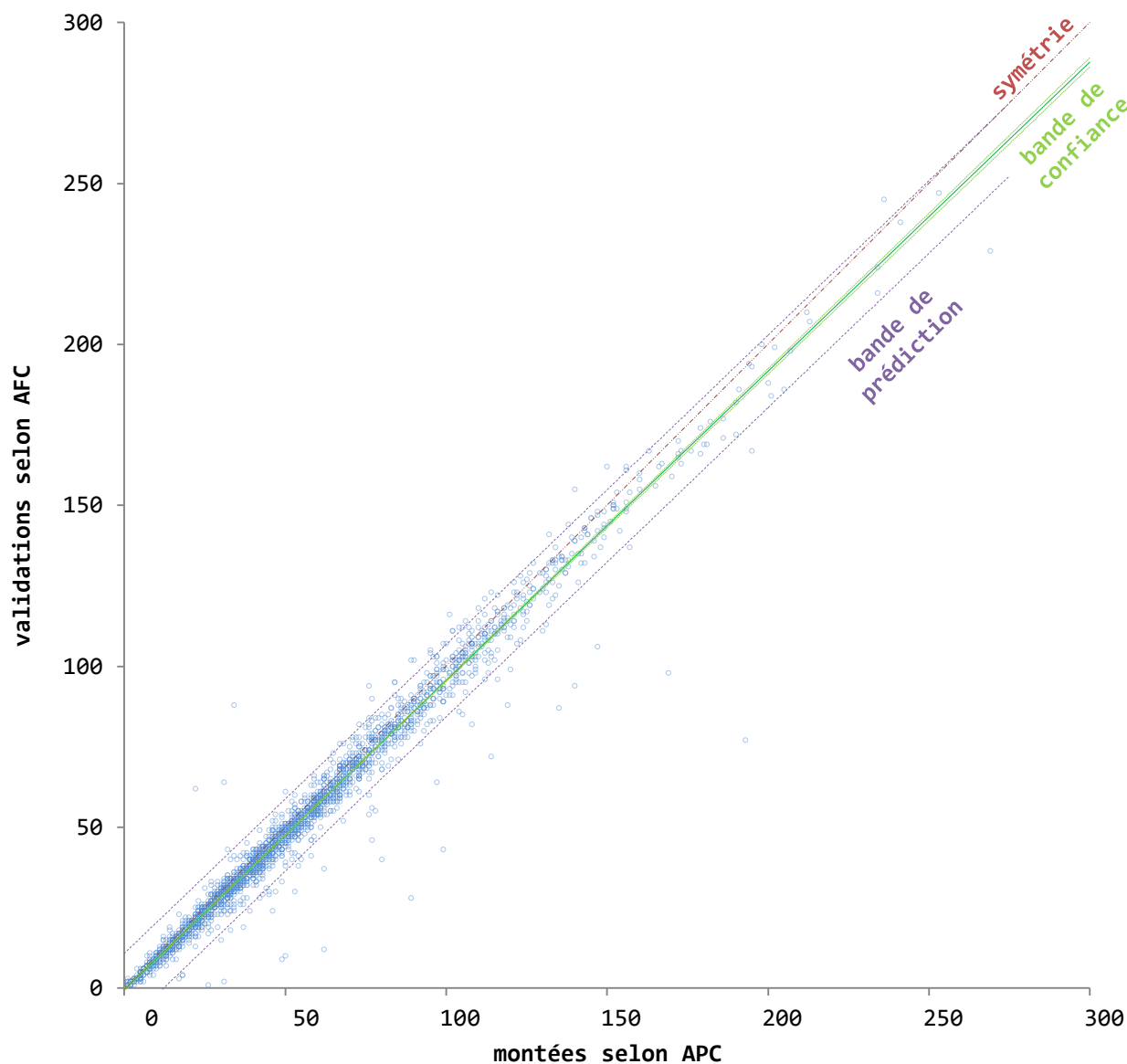
Deux modèles de régression sont construits, un confrontant les données par véhicule (figure 4-6) et l’autre par voyage (figure 4-7). Les modèles tendent à corroborer le phénomène, selon lequel les compteurs de passagers estiment plus de montants que les boîtes de perception. Puisque les modèles suggèrent que la relation entre les données des instruments est fortement linéaire, un grand volume de passagers ne semble pas expliquer la différence entre les montées comptées et les validations de titre, c’est-à-dire que la différence reste de l’ordre de 5 %, peu importe le nombre d’usagers.

La figure 4-8 montre la distribution des écarts entre les données APC et AFC en fonction du temps — les voyages avant cinq heures et après 25 heures sont à considérer avec prudence à cause de la petite taille des échantillons. En ce qui concerne la majorité des observations, l’écart ne semble pas beaucoup varier dans le temps. Par contre, les valeurs extrêmes sont particulièrement élevées aux heures de pointe, spécialement en après-midi.



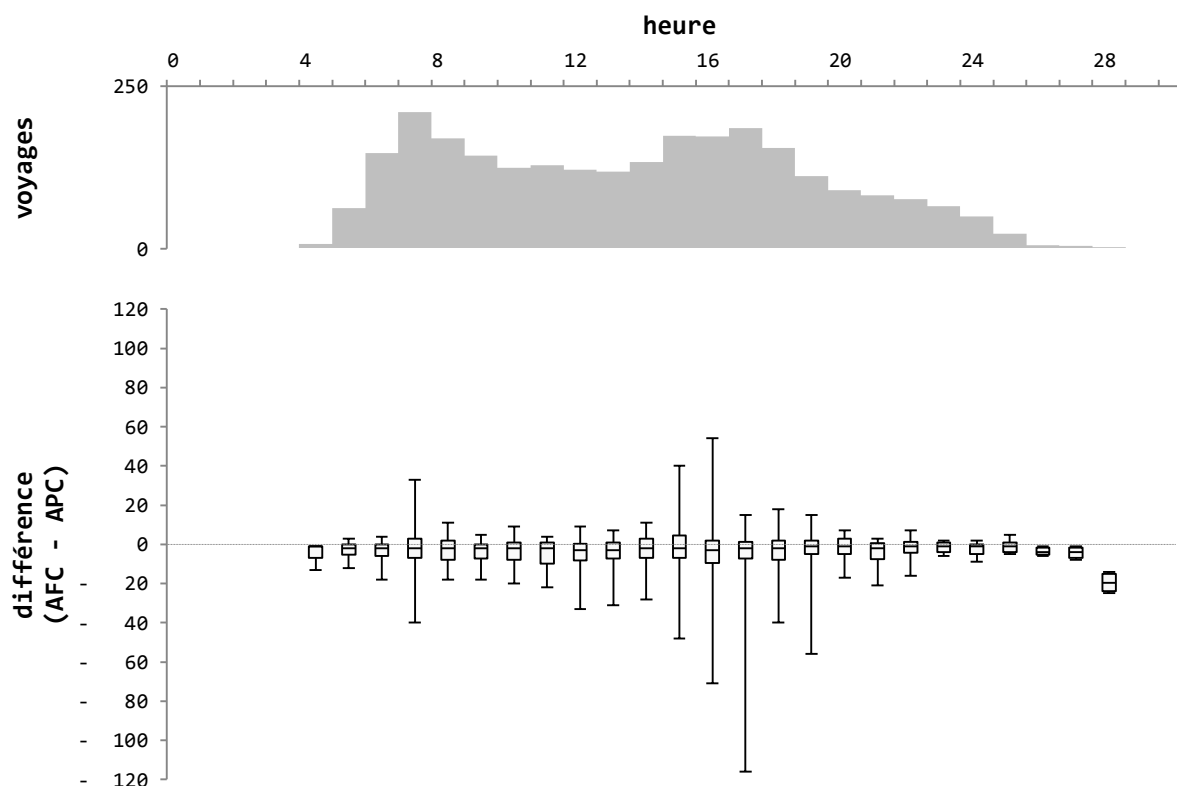
modèle de régression		$\hat{y}_i = \beta_1 x_i + \beta_0$
pente de la droite	β_1	0.95 ± 0.01
constante	β_0	4.43 ± 7.52
coefficient de détermination	R^2	0.99
taille de l'échantillon (véhicules)	n	221
coefficient de corrélation estimé	$r = \hat{\rho}$	0.997
intervalle de confiance sur ρ		$0.996 \leq \rho \leq 0.998$
seuil de confiance utilisé	$1 - \alpha$	0.95

Figure 4-6. Modèle de régression linéaire confrontant les données APC-AFC par véhicule



modèle de régression		$\hat{y}_i = \beta_1 x_i + \beta_0$
pente de la droite	β_1	0.96 ± 0.01
constante	β_0	-0.45 ± 0.40
coefficient de détermination	R^2	0.98
taille de l'échantillon (voyages)	n	2 567
coefficient de corrélation estimé	$r = \hat{\rho}$	0.988
intervalle de confiance sur ρ		$0.987 \leq \rho \leq 0.989$
seuil de confiance utilisé	$1 - \alpha$	0.95

Figure 4-7. Modèle de régression linéaire confrontant les données APC-AFC par voyage



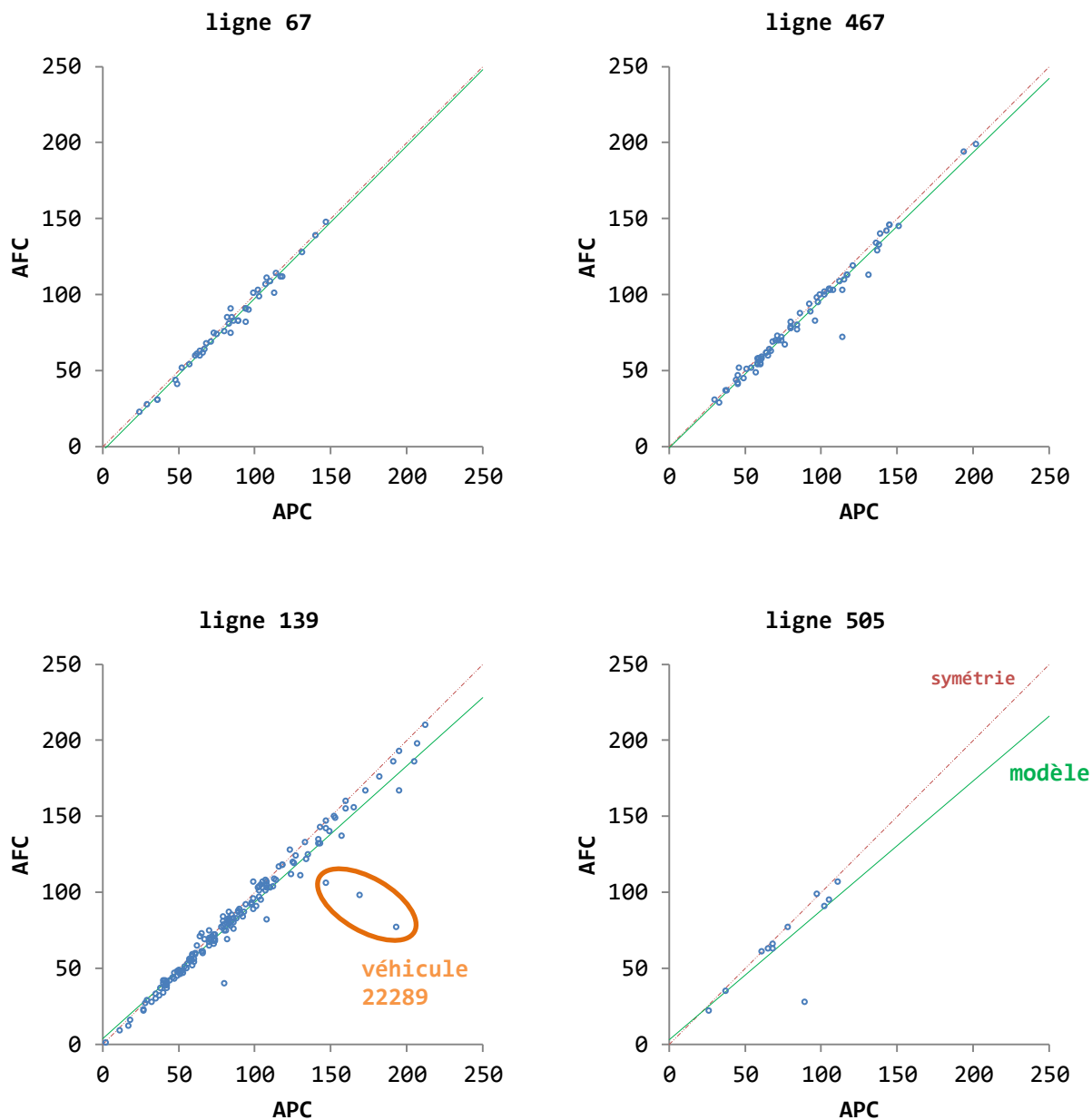
L'heure prévue de départ au premier arrêt est considérée pour chaque voyage.

Figure 4-8. Différences entre les données AFC-APC par voyage dans le temps

4.3.2 Application aux corridors

La figure 4-9 résume les modèles de régression sur les données APC-AFC par voyage pour les quatre lignes des corridors Saint-Michel et Pie-IX. Sauf pour la ligne 505 (439) à cause du faible nombre d'observations, la relation entre les deux sources de données reste très linéaire. Pour la ligne 139, plusieurs observations s'écartent loin de droite, dont trois provenant du véhicule 22289 — encadrées dans la figure 4-9.

Pour essayer d'expliquer ces valeurs extrêmes, les données de la tournée de ce véhicule, pointé aussi sur la figure 4-6, sont scrutées. Ce véhicule fait majoritairement ses voyages sur la ligne 139 : 19^e et 20^e voyages sur la 67; 27^e et 28^e voyages sur la 361; les 1^{er}, 18^e, 21^e, 26^e et 29^e voyages sont des HLP.



modèle de régression : $\hat{y}_i = \beta_1 x_i + \beta_0$
 seuil de confiance : $1 - \alpha = 0.95$

ligne		67	467	139	505
penste de la droite	β_1	1.00 ± 0.04	0.97 ± 0.04	0.90 ± 0.04	0.85 ± 0.44
constante	β_0	-2.83 ± 3.57	-0.71 ± 3.56	3.83 ± 4.00	2.79 ± 34.83
coefficient de détermination	R^2	0.98	0.97	0.92	0.65
voyages	n	44	71	160	12
coefficient de corrélation estimé	$r = \hat{\rho}$	0.992	0.987	0.961	0.809

Figure 4-9. Modèle de régression linéaire confrontant les données APC-AFC par voyage, lignes des corridors exemples

La figure 4-10 montre l'évolution du cumul des montants et des validations (le 23^e voyage n'a pas été détecté par le système AVL). Le ratio des cumuls est légèrement inférieur à l'unité au début de la tournée, soit environ 0.96 — le pic observé au tout début de la tournée est attribuable au faible nombre d'utilisateurs comptés. Le ratio diminue ensuite entre les 13^e et 17^e voyages, ou entre 15 heures et 19 heures environ, et se stabilise pour le reste de la tournée à environ 0.82. Cette période où le ratio diminue correspond à la période où les écarts AFC-APC sont importants, comme vu à la figure 4-8.

La figure 4-11 semble confirmer que plusieurs validations ne sont pas enregistrées durant la période correspondant à la pointe de l'après-midi — ou que les compteurs surestiment les montants. Cette figure semble aussi montrer que les grandes différences entre les mesures sont plutôt ponctuelles et non généralisée pendant cette période.

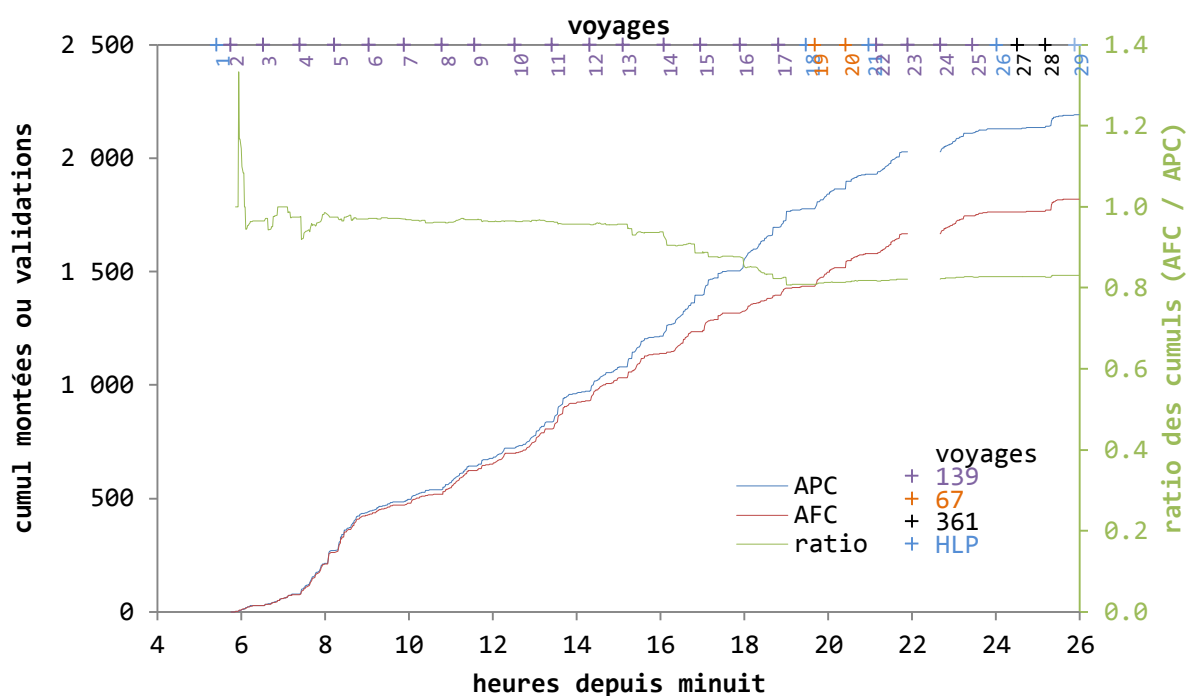


Figure 4-10. Évolution temporelle du cumul des montants, véhicule 22289

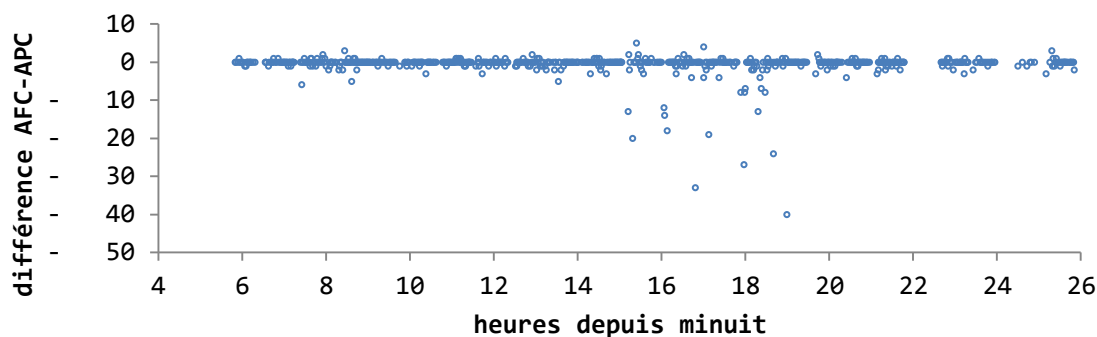


Figure 4-11. Différences par arrêt entre les données AFC et APC du véhicule 22289

D'ailleurs, le tableau 4-2 montre que ces différences sont confinées à certains arrêts; les plus grandes différences sont observées à la station de métro Pie-IX (intersection Pie-IX / Pierre-de-Coubertin). Les différences importantes observées entre les instruments pourraient être expliquées par exemple par le fait que le chauffeur ait laissé passer les usagers sans valider le passage par la boîte de perception. L'analyse des différences entre les données AFC et APC pour la ligne 139 est généralisée avec les autres véhicules à la section 4.6 par des tableaux de bord.

Tableau 4-2. Arrêts où les différences entre les données AFC et APC sont les plus grandes (véhicule 22289, ligne 139)

nom d'arrêt	tracé 1 (vers le nord)							tracé 2 (vers le sud)						
	n°Séq	nombre	centiles des différences (AFC - APC)					n°Séq	nombre	centiles des différences (AFC - APC)				
			0	25	50	75	100			0	25	50	75	100
Ontario	6	8	-13	-2	0	0	0	40	10	0	0	0	0	0
Pierre-de-Coubertin	10	9	-40	-19	-1	0	3	36	10	-33	-3	-2	0	0
Jarry	27	9	-3	-1	0	0	1	20	10	-13	-2	0	0	0
Gariépy / Henri-Bourassa								2	9	-27	-6	0	0	0
Garon / Henri-Bourassa	47	8	-12	0	0	0	0							
Henri-Bourassa / Pie-IX	48	-	-	-	-	-	-	1	9	-14	-3	-1	0	0

*n°Séq : numéro de l'arrêt dans la séquence / nombre : nombre de passages détectés (voir annexe 4 pour le tableau complet)

En somme, cette confrontation révèle les défis d'intégration de sources de données potentiellement complémentaires et en apparence redondantes, dans le sens où les systèmes APC et APC permettent tous deux d'estimer le nombre de montants. Les différences révélées permettent au planificateur d'entreprendre des analyses plus poussées sur leurs causes et de prendre les actions nécessaires pour corriger la situation.

4.4 Ponctualité en début et en fin de voyage

Connaissant les heures de passage aux arrêts, autant « planifiées » — ou affichées pour l’usager — que « réelles », il est possible d’évaluer la ponctualité des autobus aux différents arrêts en effectuant une simple différence entre ces temps. Cette section aborde cet indicateur de manière globale en étudiant les débuts et les fins de voyages — il serait en effet plus difficile de comparer les autres arrêts entre plusieurs lignes différentes. La section 4.6 montre la ponctualité des autobus pour chaque arrêt de certaines lignes.

La Société de transport de Montréal (2012b) s’est donné une norme de ponctualité selon laquelle les autobus devraient passer à un arrêt entre une minute avant et trois minutes après l’heure planifiée. Cette norme est réutilisée dans ce mémoire pour donner des points de repère lors de l’évaluation de la ponctualité des autobus. Pour ce faire, la distribution de la différence entre les heures réelle et planifiée est utilisée. Cette distribution est illustrée par des boîtes à moustaches (voir section 4.2).

La ponctualité en début et en fin de voyage — départ au premier arrêt et arrivée au dernier arrêt respectivement — est montrée aux figures 4-12 et 4-13 par boîtes à moustaches et par courbes de répartition respectivement. La figure 4-14 illustre la ponctualité en début et en fin de voyage pour les voyages productifs selon l’heure prévue de départ au premier arrêt. Il faut noter que les données avant cinq heures et après 25 heures sont à considérer avec prudence à cause du faible nombre de voyages.

Les autobus en voyage productif semblent débiter à l’heure et tendent à être légèrement en avance au dernier arrêt. Toutefois, il faut comprendre que les points de contrôle en fin de ligne productive sont moins importants que les précédents dans le respect de l’horaire prévu du point de vue des usagers. En effet, il est raisonnable de penser qu’arriver en avance au terminus d’arrivée ne devrait pas poser problème pour la plupart des usagers.

Quant aux interlignes et aux voyages d’injection et d’éjection, les autobus semblent majoritairement arriver en avance sur le temps prévu. Cela pourrait remettre en question la planification et l’organisation des tournées, surtout pour les haut-le-pied depuis ou vers les

garages. Il ne faut toutefois pas oublier que le temps alloué aux interlignes peut aussi inclure le temps de repos pour le chauffeur ou le temps de recouvrement servant à compenser le retard accumulé au voyage précédent.

Cette section montre donc ce qui se passe au début et à la fin des voyages, ce qui est très pertinent pour les voyages non productives, puisqu'ils n'ont que deux arrêts normalement.

L'évolution du service à l'intérieur des voyages productifs pour les corridors exemples des boulevards Saint-Michel et Pie-IX est vue avec les diagrammes espace-temps et avec les tableaux de bord dans les prochaines sections de ce chapitre.

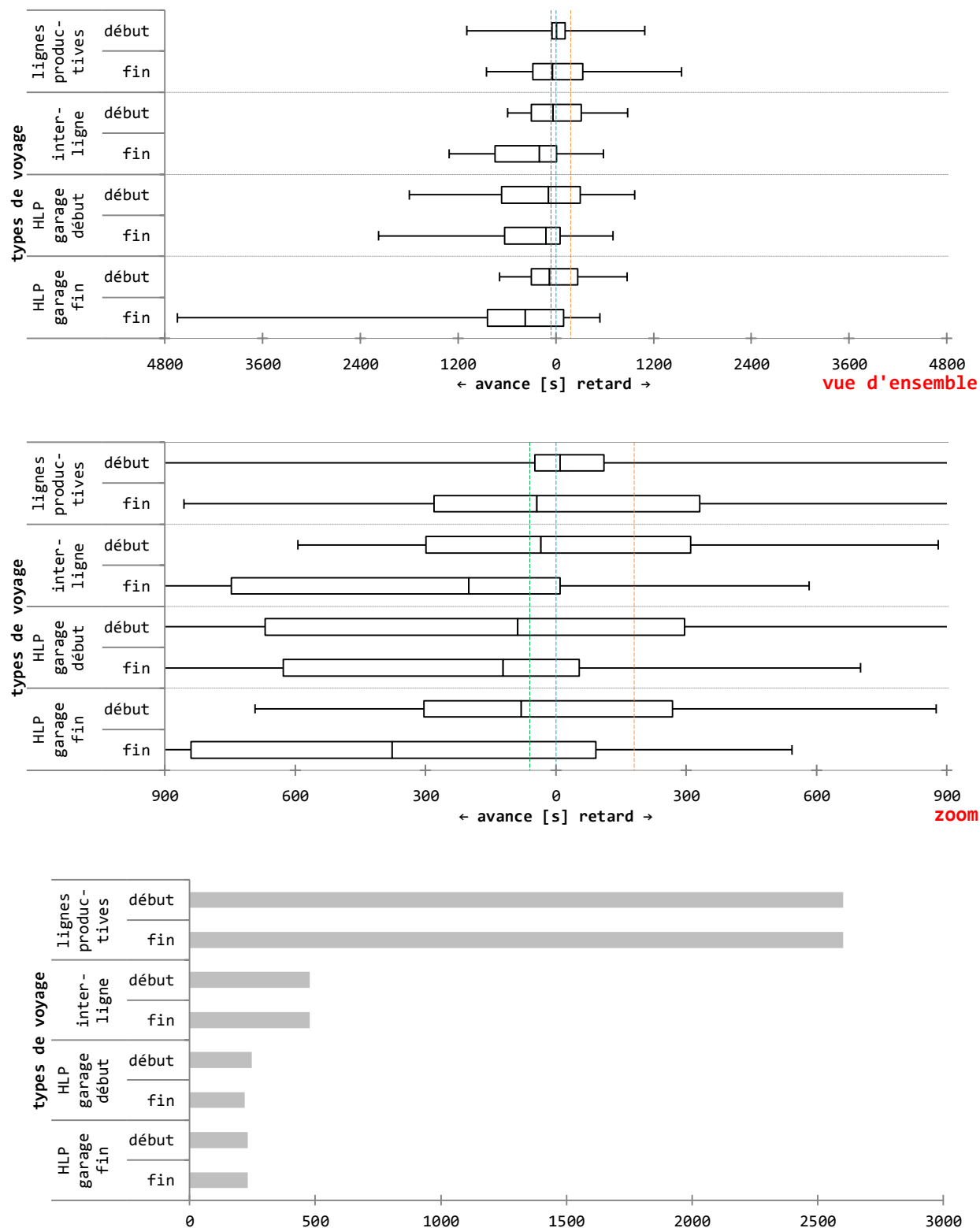


Figure 4-12. Ponctualité en début et en fin de voyage selon le type de voyage

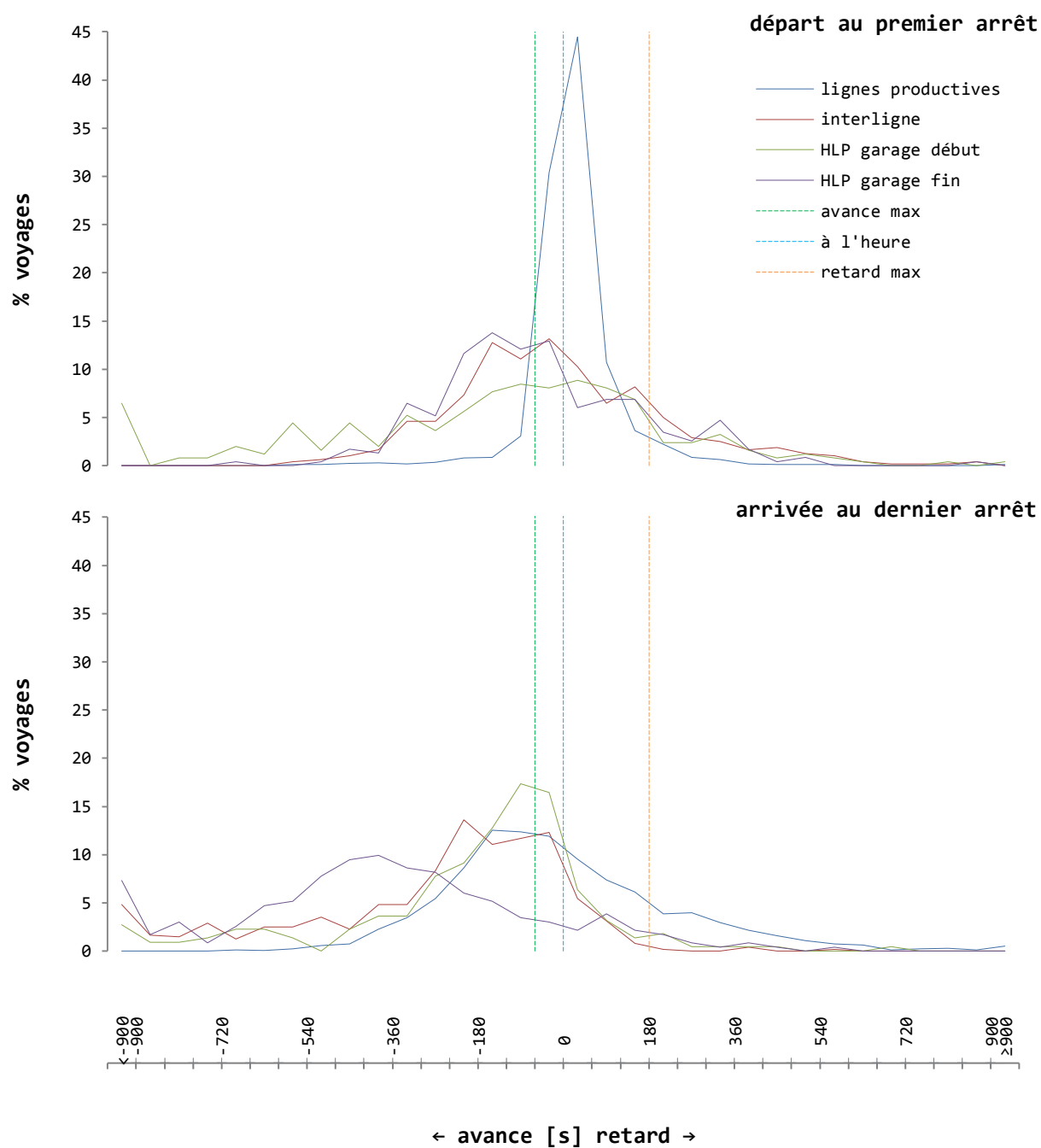


Figure 4-13. Répartition de la différence entre les heures de passage réelles et planifiées en début et en fin de voyage

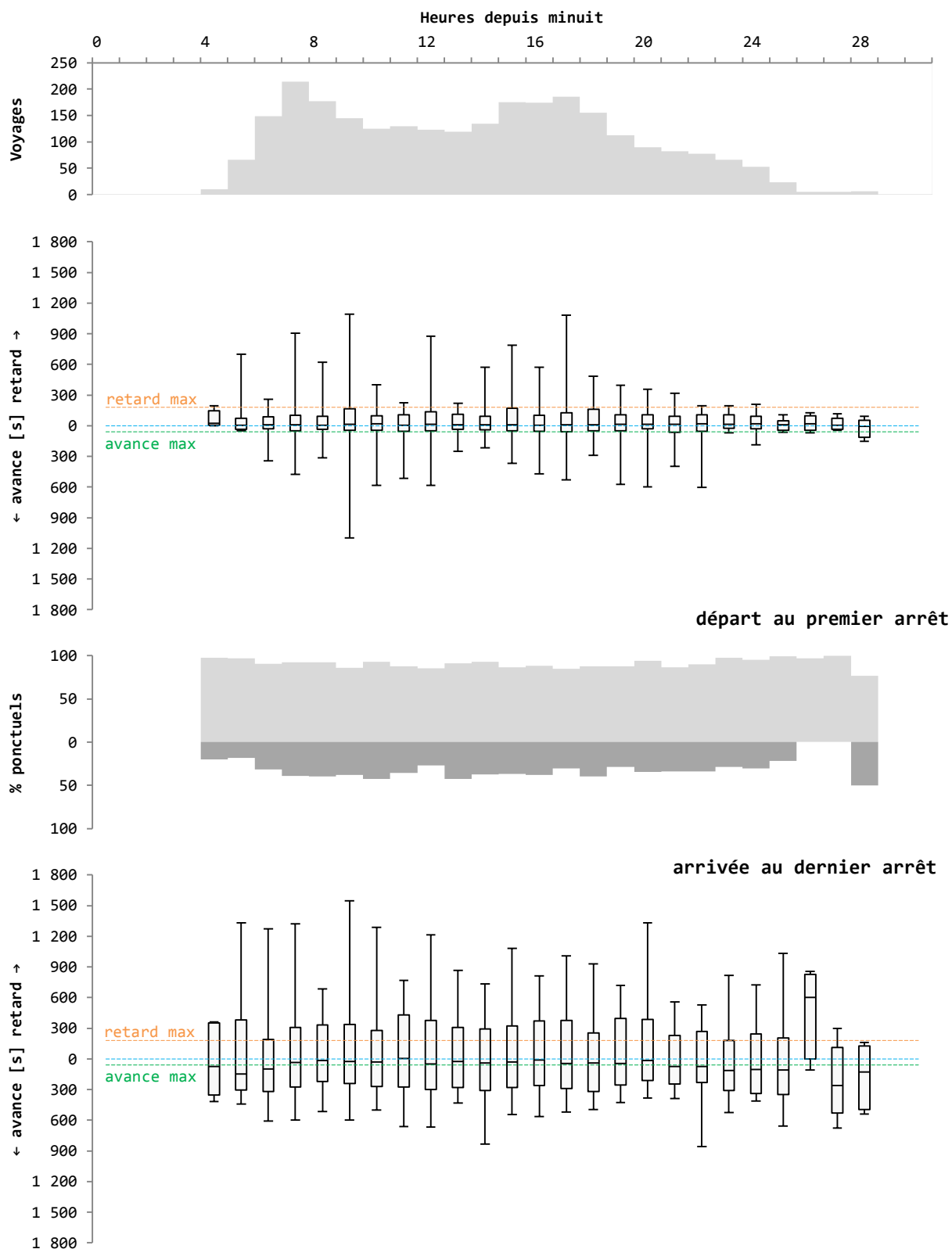


Figure 4-14. Ponctualité des lignes productives en début et en fin de voyage

4.5 Diagramme distance-temps interactif

Un diagramme espace-temps, ou distance-temps (Vuchic, 2007), est un graphique de courbes mettant la distance parcourue par un ou plusieurs véhicules en fonction du temps. Le diagramme devient un outil de visualisation particulièrement intéressant en permettant, notamment, une analyse du parcours des autobus d'arrêt en arrêt ou en l'utilisant comme moyen de conception des lignes de transport (Vuchic, 2007). Alors que la pente des courbes permet d'estimer la vitesse du véhicule, le diagramme peut aussi mettre en lumière les dépassements ou les formations de pelotons entre plusieurs véhicules.

D'autres éléments peuvent aussi s'ajouter au diagramme espace-temps, comme la charge des autobus ou les mouvements d'usagers aux arrêts. Toutefois, afficher tout cela pourrait surcharger le diagramme et le rendre ainsi illisible. L'idée présentée dans cette section est d'ajouter à l'utilisateur la possibilité d'interagir avec le diagramme en lui laissant le choix d'afficher ou de masquer ou de modifier rapidement les éléments du graphique.

4.5.1 Construction et description du diagramme

La figure 4-15 résume les fichiers entrants nécessaires et la construction du diagramme distance-temps. Les arrêts d'une ligne sont préalablement linéarisés et distancés par rapport au précédent en fonction de la médiane des distances interarrêts trouvées avec le système AVL. Les arrêts communs aux lignes-directions — à la même intersection — sont mis ensemble (fichier des arrêts). Chaque voyage est ensuite représenté sur le diagramme. Pour ce faire, les données sont converties en fichier texte (.csv) comprenant les temps de passage et les mouvements de passagers aux arrêts pour chaque voyage (fichier passages).

Ensuite, ces données sont traitées à l'aide d'un code JavaScript pour les représenter sous forme de diagramme en format vectoriel (SVG). Ce code utilise la bibliothèque graphique JavaScript D3.js (Bostock, 2012) pour dessiner les différents éléments du diagramme. En combinaison avec une interface HTML/CSS, le diagramme permet d'afficher, de cacher ou de modifier certains éléments du diagramme. Des détails supplémentaires sur le fonctionnement du code JavaScript servant à générer les diagrammes se trouvent en annexe 5. De plus, des extraits de fichiers pour le corridor Saint-Michel sont montrés en annexe 6.

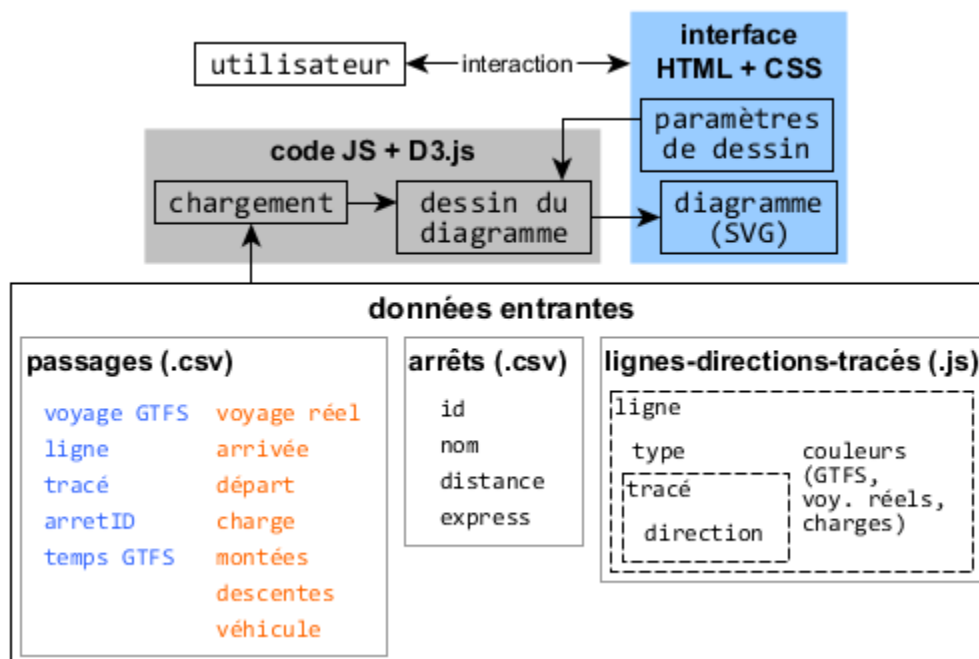


Figure 4-15. Données entrantes et génération du diagramme distance-temps

La figure 4-16 montre à titre d'exemple un extrait du diagramme distance-temps du corridor Saint-Michel en direction sud. Les arrêts présents dans la figure indiquent que la plupart des arrêts entre Prieur et la station de métro Saint-Michel sont communs à toutes les lignes du corridor. Les arrêts avant Prieur sont propres à certaines lignes-tracés.

Les tracés de la ligne 505 sont différents de ceux de la ligne 139 dans la partie nord, ce qui peut conduire à un chevauchement des arrêts sur le diagramme, comme dans l'exemple illustré à la figure 4-19.

Les voyages GTFS (couleur pâle) et « réels » (couleur foncée) sont ainsi construits, sachant l'emplacement des arrêts sur le diagramme et les temps de passage. Les courbes des voyages « réels » sont accompagnées de traits épaissis proportionnellement à la charge du véhicule. Les éléments du diagramme en mauve représentent les voyages des lignes locales (67 et 139); en bleu, ceux des lignes express (467 et 505/439). À cause des particularités de la détection par GPS (voir section 3.2.2), l'arrivée au premier arrêt et le départ au dernier arrêt ne sont pas représentés dans les diagrammes.

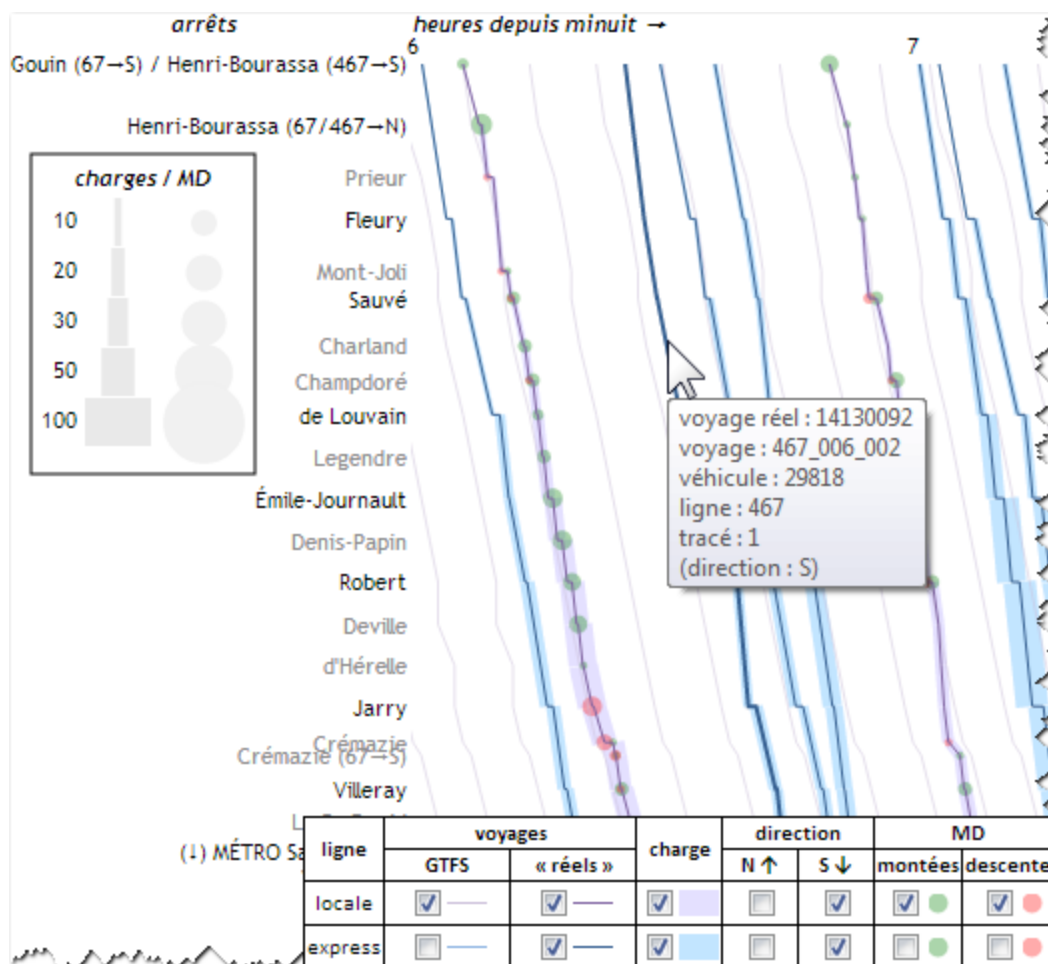


Figure 4-16. Extrait de diagramme distance-temps et exemple d'infobulle, corridor Saint-Michel

Les montées et les descentes — du système APC — sont ajoutées à chaque arrêt par des disques dont l'aire est proportionnelle au nombre d'utilisateurs. Les descentes sont placées à l'arrivée de l'autobus à l'arrêt et sont rouges, alors que les montées, en vert, sont mises au départ de l'arrêt, principalement pour limiter la superposition des disques sur le diagramme.

L'utilisateur qui se sert du diagramme distance-temps peut choisir d'afficher ou de masquer certains éléments ou de modifier leurs caractéristiques, comme la couleur, l'échelle des charges ou des montées-descentes (MD) ou la taille du texte (voir annexe 7). L'utilisateur peut aussi scruter à la loupe les détails du diagramme : en pointant sur un objet, une infobulle apparaît avec des informations supplémentaires, comme vu à la figure 4-16. De plus, lorsqu'un voyage est pointé, ses courbes GTFS et « réelles » sont épaissies légèrement et sont plus foncées. Aussi, l'utilisateur peut utiliser les outils de son navigateur Web pour examiner les éléments plus en détail, comme le zoom.

4.5.2 Application aux corridors

Comme mentionné plus haut, l'utilisateur est submergé d'informations si tous les éléments du diagramme sont affichés. Cette section montre donc des points importants du diagramme distance-temps, illustrés par les corridors exemples Saint-Michel et Pie-IX, en faisant varier l'affichage ou le cachement de certains éléments.

La figure 4-17 montre un extrait du diagramme espace-temps pour la ligne 67 sud, entre 04:30 et 07:00 en n'affichant que les voyages GTFS et « réels ». La résolution temporelle des voyages GTFS à la minute introduit quelques aberrations, car la courbe devient verticale entre deux arrêts si les temps de passage sont les mêmes. Cela n'empêche toutefois pas de comparer les voyages GTFS aux voyages « réels » pour avoir une idée de la conformité à l'horaire des véhicules. Par exemple, le premier voyage « réel » visible dans le diagramme de la figure 4-17 a pris de plus en plus d'avance par rapport à son horaire.

La figure 4-17 comporte des « trous » entre les voyages « réels », c'est-à-dire qu'il manque plusieurs voyages. D'ailleurs, le diagramme de la ligne 467 sud, à la figure 4-18, montre très bien l'étendue de ces « trous » pour cette ligne-direction. Le fait qu'une partie de la flotte d'autobus sont instrumentées explique largement ce phénomène.

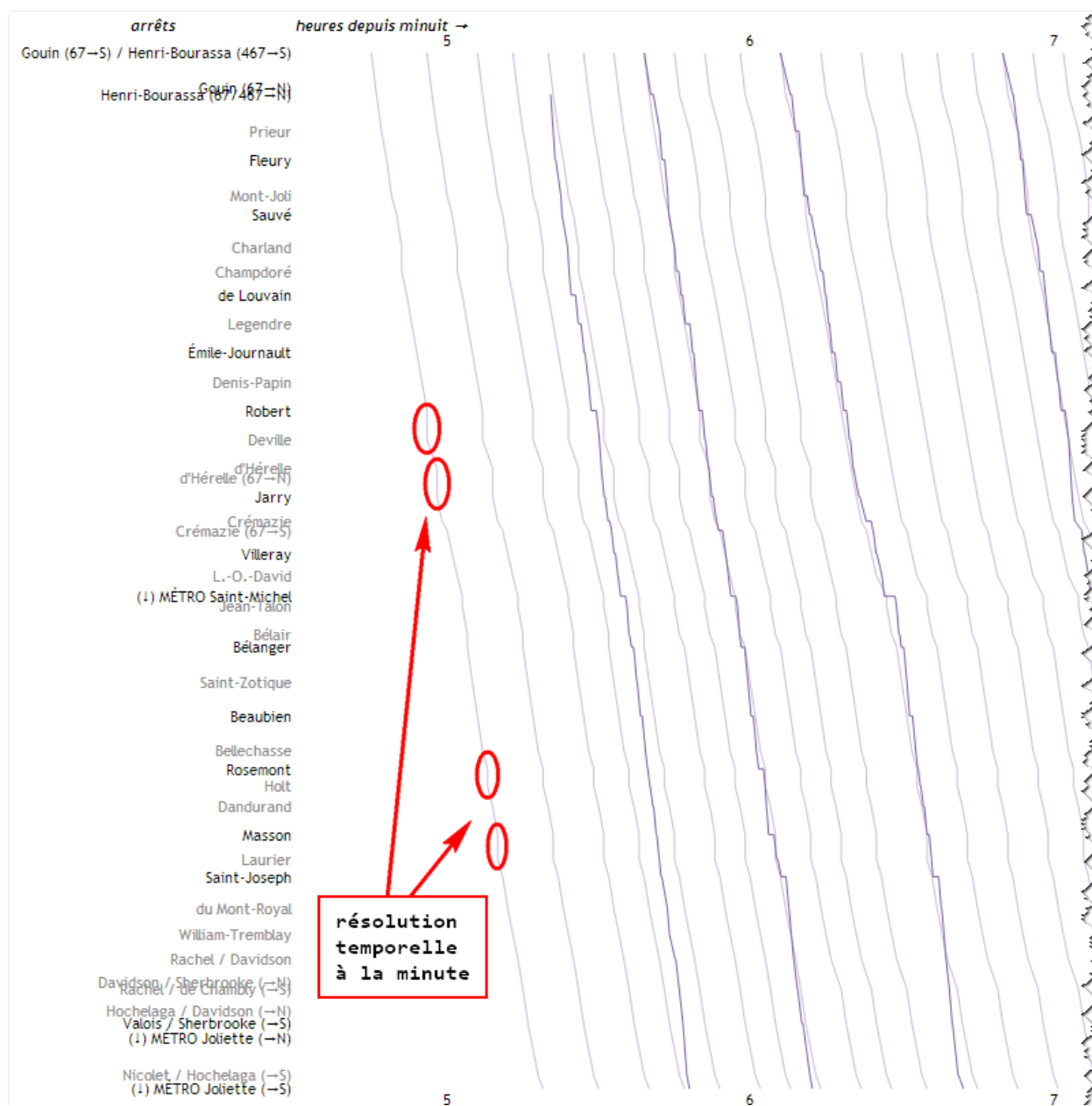


Figure 4-17. Extrait de diagramme distance-temps, ligne 67 sud, voyages GTFS et « réels »

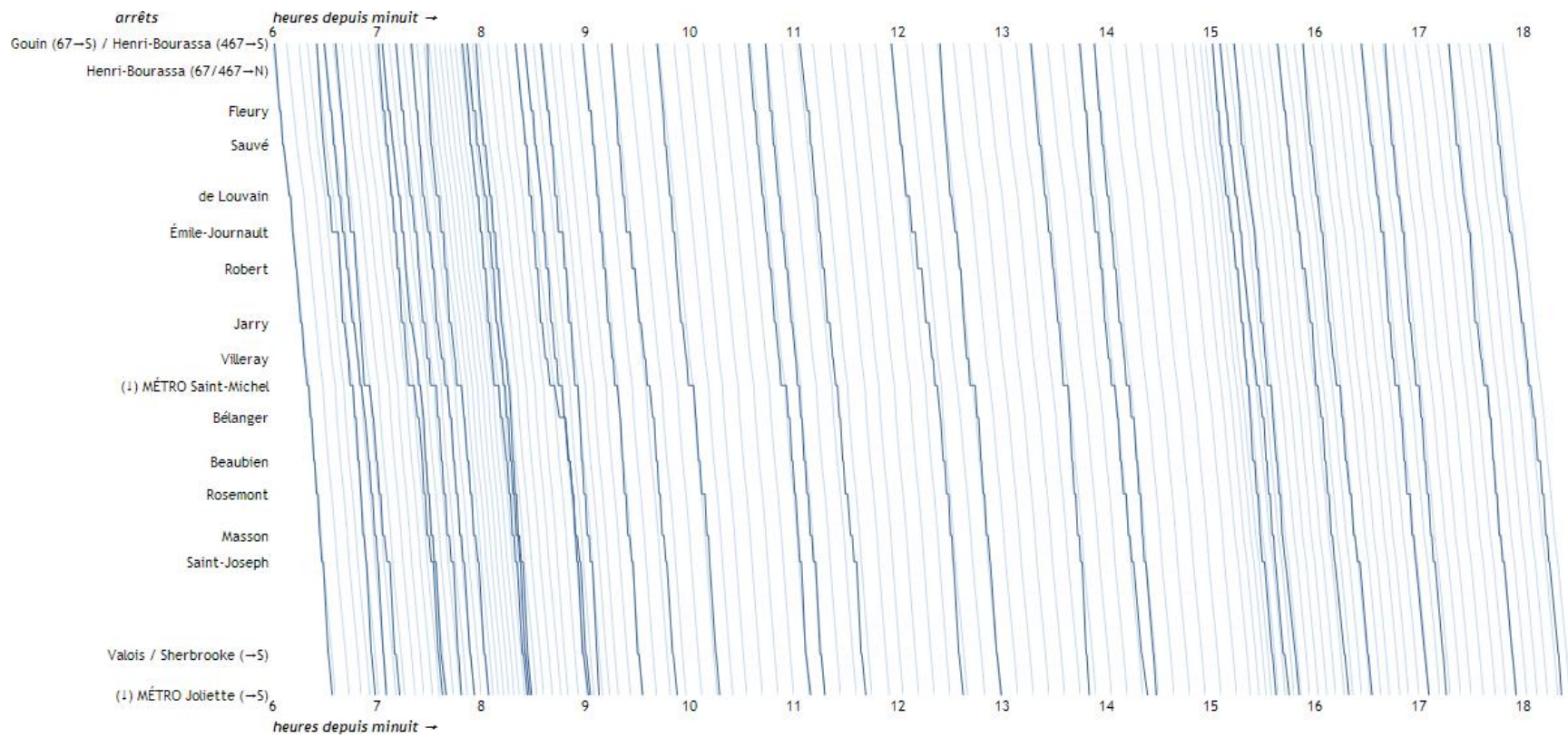


Figure 4-18. Diagramme distance-temps, ligne 467 sud

La figure 4-19 illustre un cas intéressant de ce qui semble être de la congestion sur le boulevard Pie-IX en période de pointe de l'après-midi en direction nord :

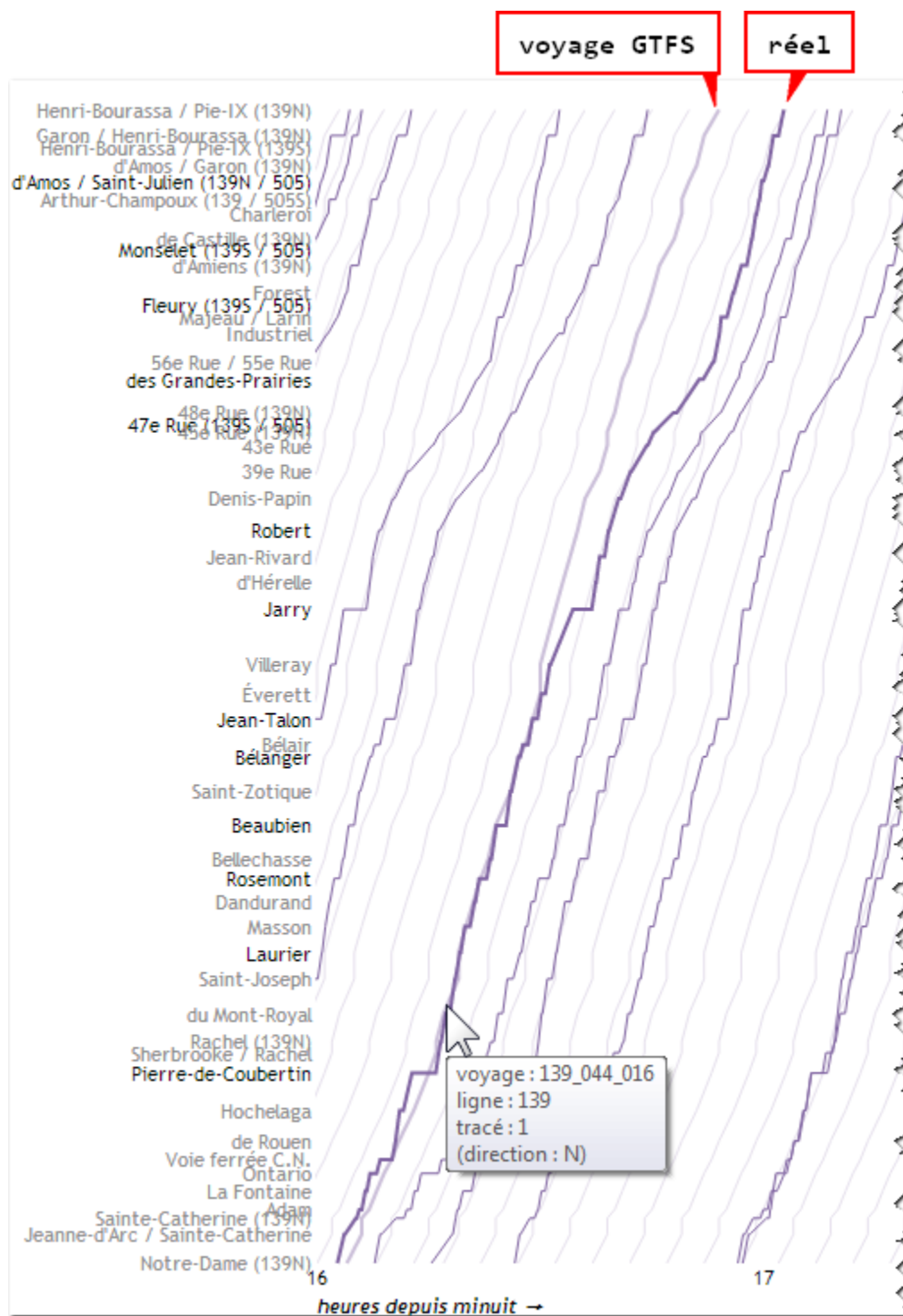


Figure 4-19. Diagramme distance-temps, voyage 139_004_016, ligne 139 nord

L'ensemble des autobus sur la figure ralentissent clairement entre les boulevards Robert et des Grandes-Prairies. La figure pointe le voyage 139_044_016 du véhicule 22296 et permet de constater l'ampleur des répercussions de ce ralentissement du service sur la ponctualité : le décalage entre la courbe GTFS et « réel » atteint plus de dix minutes dans ce cas. La forme similaire des courbes des voyages « réels » dans la période présente dans la figure suggère qu'un même phénomène explique ces ralentissements.

La figure 4-20 montre deux pelotons d'autobus sur le corridor Saint-Michel en direction sud entre huit heures et neuf heures. Un peloton se forme, souvent involontairement, lorsque deux ou plusieurs autobus se trouvent au même arrêt ou endroit, alors qu'il devrait y avoir un espace entre eux selon l'horaire planifié, ce qui peut se produire lorsqu'au moins un de ces autobus a pris trop de retard ou d'avance ("Bus bunching," s.d.; Transportation Research Board & Kittelson and Associates Incorporated, 2003). Ce phénomène est encore plus particulier avec des autobus d'un même corridor mais de lignes différentes.

Dans cette figure 4-20, seuls les voyages GTFS correspondant aux voyages « réels » sont affichés. Le premier pelotonnement, des véhicules 29820 et 29822, survient à l'intersection avec la rue Masson, mais reste bref ou ponctuel malgré l'avance du véhicule 29822. Par contre, la formation du deuxième peloton, entre les véhicules 29819, 26085 et 29815, est plus complexe. Le véhicule 29819 desservant la ligne 467 prend du retard après la station Saint-Michel et forme un peloton avec le véhicule 29815, qui circule pour la ligne 467 aussi, jusqu'à la rue Masson. Le véhicule 26085, sur la ligne 67, a pris du retard dès le début du trajet et se fait rattraper par le véhicule 29815, alors que ce dernier ne devait le faire qu'à l'approche de l'intersection des rues Rachel et Davidson, vers la fin du voyage.

En ajoutant les charges et les montées des véhicules, ce qui donne la figure 4-21, le diagramme donne un aperçu du comportement des usagers face à ce phénomène. Dans le premier peloton, le véhicule 29822, étant en avance après la station Saint-Michel et suivant le véhicule 29820, ne fait monter que peu d'usagers. Dans le deuxième peloton, le véhicule 29815, entre les deux autres véhicules du peloton, ramassent plus d'usagers que les deux autres véhicules, le premier étant très en retard et le dernier faisant monter les usagers pour les arrêts ne faisant pas partie du service express.

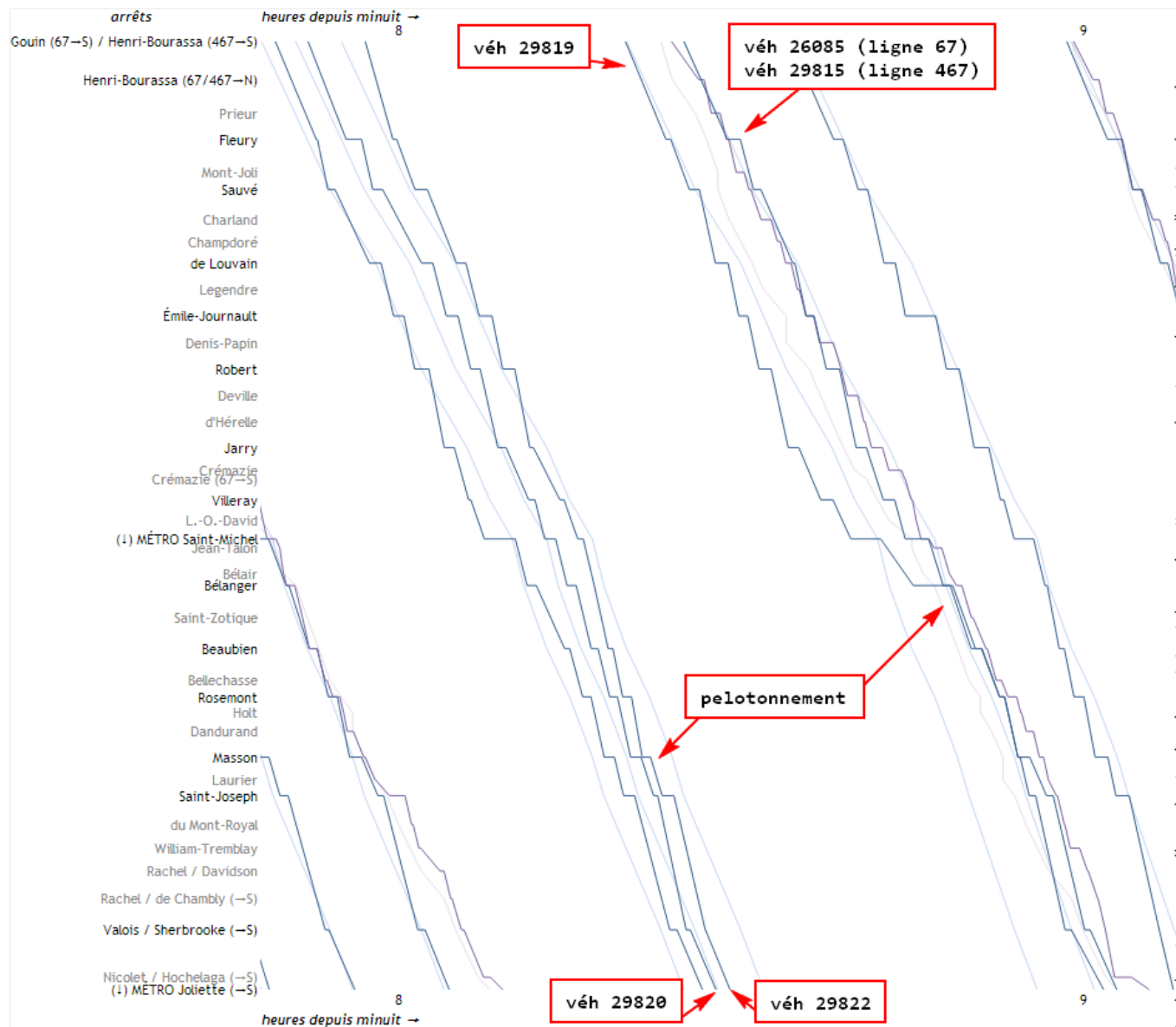


Figure 4-20. Diagramme distance-temps, corridor Saint-Michel, direction sud, entre 08:00 et 09:00

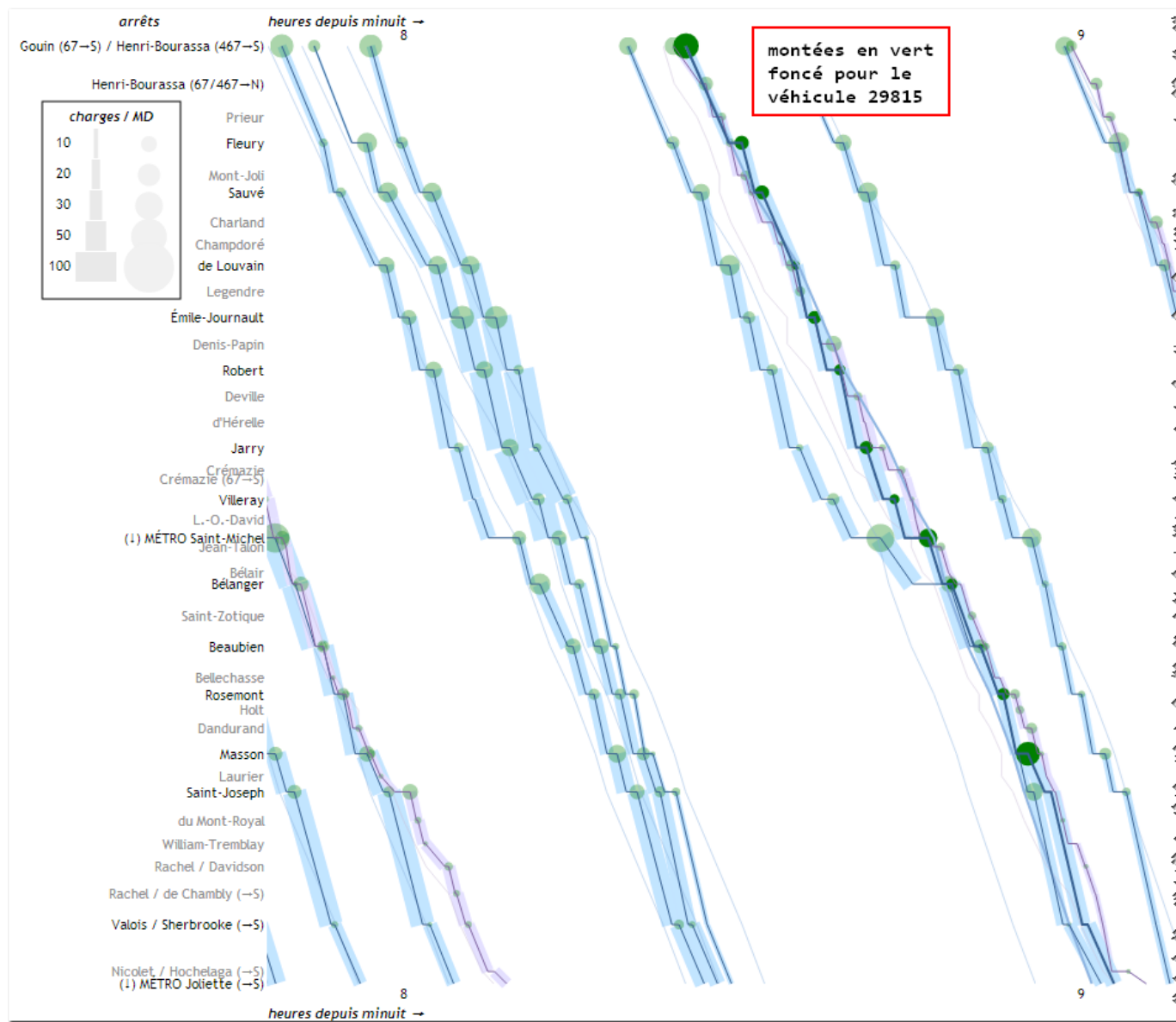


Figure 4-21. Diagramme distance-temps, corridor Saint-Michel, direction sud, entre 08:00 et 09:00, avec charges

La figure 4-22 illustre les voyages du corridor Pie-IX en direction nord entre 15:00 et 18:00 avec leur charge, les montées et les descentes aux arrêts. Cet exemple montre bien que le diagramme espace-temps interactif peut faire dégager des tendances, non seulement pour les temps de parcours, comme vu précédemment pour la congestion, mais aussi dans les mouvements de passagers.

Dans le cas du corridor Pie-IX en direction nord, les montées semblent concentrées dans la partie sud et à quelques arrêts seulement au nord, comme aux rues Jean-Talon et Jarry. Les descentes sont quant à elles relativement réparties tout au long du trajet au nord de la station Pie-IX (Pierre-de-Coubertin). Cette répartition des mouvements suggère donc que le corridor sert surtout à amener les gens de la station de métro vers Montréal-Nord à cette période. D'ailleurs, les tableaux de bord, à la section 4.6, permettent de corroborer cette tendance.

En somme, le diagramme distance-temps interactif permet au planificateur de scruter à la loupe les voyages d'un corridor en jouant avec les propriétés des objets de la figure. En profitant des propriétés du diagramme, soit la linéarisation des arrêts et la transposition temporelle de la distance parcourue par les véhicules, l'utilisateur peut aussi afficher la charge et les mouvements de passagers. Avec les voyages GTFS et « réels », le planificateur peut évaluer la ponctualité et la conformité par rapport à l'horaire et observer les perturbations de service. Le diagramme distance-temps donne ainsi l'opportunité de mieux décrire le lien entre le comportement des usagers et l'utilisation du service.

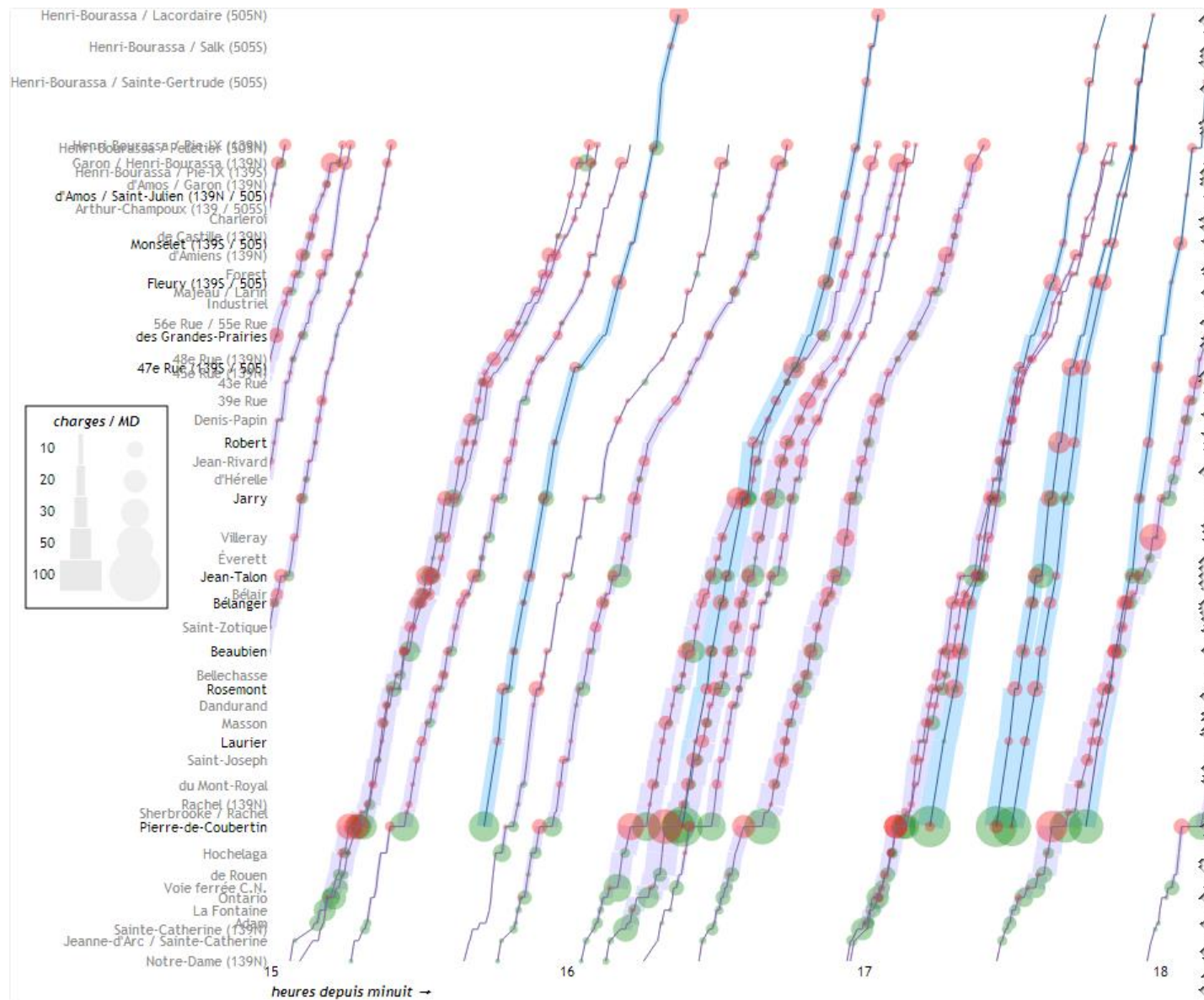


Figure 4-22. Diagramme distance-temps, corridor Pie-IX, direction nord, avec charges, montées et descentes

4.6 Tableau de bord

Un tableau de bord présente sommairement et de manière ciblée les indicateurs pertinents et essentiels d'un système; l'agencement des indicateurs permet de montrer une vision globale du système et met en évidence les résultats significatifs et les tendances, enrichissant ainsi l'analyse et l'interprétation de l'information et appuyant le gestionnaire dans le suivi du système, dans le diagnostic de problèmes ou dans la prise de décisions (Voyer, 2000). Le tableau de bord, comme ceux dans les voitures, les autobus ou les cockpits des avions, devient alors un outil efficace de visualisation et de communication de l'information en la synthétisant.

4.6.1 Construction et description du tableau de bord

Le tableau de bord proposé montre divers indicateurs pour les lignes-tracés. Pour chaque arrêt, le lecteur peut lire dans l'ordre le nombre de passages détectés, les distances interarrêts, la ponctualité, la durée de la halte aux arrêts, les informations sur les montées et sur les descentes et le cumul de la charge et des montants. Ces indicateurs sont précalculés et convertis en fichier .csv compatible pour le code JavaScript qui dessine le tableau de bord — comme pour les diagrammes distance-temps, ce code utilise la bibliothèque D3.js pour exporter le tableau en dessin vectoriel SVG (voir section 4.5.1). La figure 4-23 résume les données entrantes et la construction du tableau de bord. Un extrait du fichier entrant est montré en annexe 8.

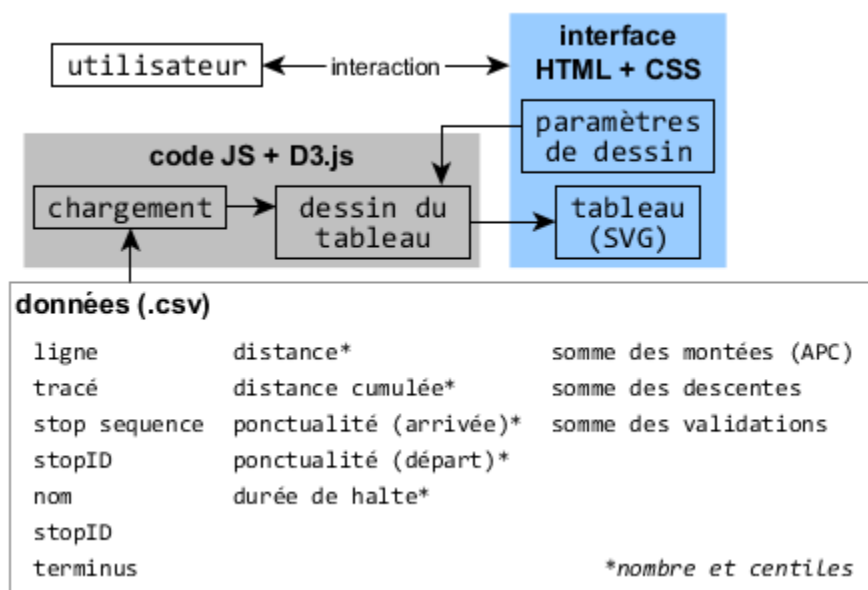


Figure 4-23. Données entrantes et génération du tableau de bord

Comme pour le diagramme distance-temps, les divers objets du tableau de bord peuvent être inspectés en les pointant, comme le montre la figure 4-24. Les différents objets du tableau de bord sont personnalisables. Les différents paramètres sont présentés en annexe 7.

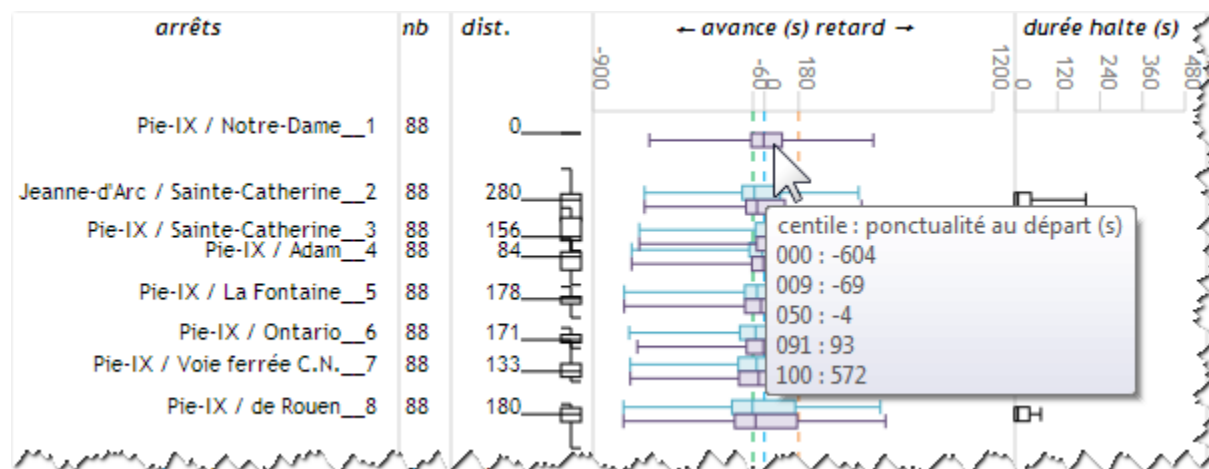


Figure 4-24. Exemple d'infobulle pour les objets du tableau de bord

La figure 4-25 montre le tableau de bord pour la ligne 467 en direction sud; la figure 4-26, en direction nord. Les figures 4-27 et 4-28 présente le tableau de bord pour la ligne 67, respectivement pour les directions nord et sud. Quant au corridor Pie-IX, les tableaux de bord pour la ligne 439/505 sont aux figures 4-29 (nord), 4-30 (sud, tracé 2) et 4-31 (sud, tracé 3) et ceux pour la ligne 139 aux figures 4-32 (nord) et 4-33 (sud). Étant donné la grande taille des figures, elles sont placées à la fin de ce chapitre. De plus, la figure 4-25 contient des précisions sur certains éléments du tableau qui sont présents aussi dans les autres figures.

D'abord, le nombre de passages détectés est l'équivalent de la taille de l'échantillon et indique au lecteur si les valeurs des indicateurs sont représentatives de ce qui se passe à l'arrêt. Chaque arrêt est distancé par rapport au précédent proportionnellement à la médiane de la distance interarrêts. La ligne-tracé est ainsi linéarisée. Ensuite, les distances interarrêts sont montrées en valeur numérique et sous forme de boîte à moustaches. Plus la boîte est petite, moins il y a de risques qu'il y ait un problème de détection à cet arrêt. À l'inverse, si les moustaches ou les boîtes de deux arrêts se chevauchent, cela peut indiquer que les arrêts sont très proches ou que le système n'arrive pas à les discriminer. De grandes moustaches peuvent indiquer des valeurs extrêmes isolées (voir section 4.2 pour les centiles utilisés). Si un arrêt est systématiquement non

détecté, l'arrêt suivant est placé en fonction de la distance depuis le premier arrêt — d'où le nom de l'arrêt en rouge et l'abréviation « N/A » sur la figure 4-25 pour indiquer à l'utilisateur qu'il y a un problème de détection de l'arrêt précédent.

La ponctualité des autobus aux arrêts est indiquée par des boîtes à moustaches bleu-vert pour l'arrivée et mauve pour le départ. Sauf au dernier arrêt, la ponctualité au départ de l'arrêt est généralement celle qui intéresse le plus — l'usager peut monter dans l'autobus ou descendre tant que le véhicule est arrêté. En raison des particularités concernant la détection du premier et du dernier arrêt (voir section 3.2), la boîte à moustaches d'arrivée au premier arrêt et celle de départ au dernier arrêt sont omises. Des lignes repères indiquant les normes de ponctualité (voir section 4.4) montrent dans quelle mesure les autobus respectent l'horaire.

Le décalage entre les boîtes à moustaches de départ et d'arrivée donne une indication, une estimation de la durée de la halte. Des boîtes à moustaches en noir et blanc pour la durée de la halte aux arrêts clarifient leur distribution.

Ensuite, quatre diagrammes à bandes donnent de l'information sur les montants et sur les descendants par les systèmes AFC et APC. Le premier diagramme indique la différence entre le nombre de validations de passages selon le système AFC et les montées comptées par l'APC à la porte avant de l'autobus. Le deuxième diagramme indique le cumul des validations de passages; le troisième, le cumul des comptages de montées, peu importe la porte d'autobus; le quatrième, le cumul des descentes. Ces diagrammes à bandes permettent de diagnostiquer les problèmes avec ces systèmes. Par exemple, pour la ligne 467 (figures 4-25 et 4-26), la différence entre les systèmes AFC et APC semble très élevée au premier arrêt et à la station Saint-Michel — il est à noter que l'échelle du premier diagramme n'est pas la même que les trois autres. Ces diagrammes permettent aussi de faire ressortir rapidement les arrêts les plus achalandés.

Le cumul des charges est la somme des charges de tous les voyages de la ligne-tracé. Pour chaque arrêt, la somme des montées, selon les compteurs automatiques, est ajoutée à la charge de l'arrêt précédent. Ensuite, la somme des descentes est soustraite pour obtenir le cumul de la charge pour cet arrêt. Cela permet d'estimer l'utilisation de la ligne et de voir en un coup d'œil les points critiques de la ligne — là où les charges sont les plus élevées. Au dernier arrêt de la ligne, l'utilisateur peut voir l'ampleur du débalancement des montées et des descentes fournies par le système APC : plus la courbe est près de zéro, moins le débalancement est important. Par contre,

s'il manque des voyages à certains arrêts, l'évaluation de la différence montées-descentes APC est à utiliser avec prudence.

Enfin, la courbe du cumul des montants compare l'évolution du cumul des montées APC à celle des validations de passages du système AFC. Cette courbe permet de noter notamment la propagation de la différence entre les mesures deux systèmes.

4.6.2 Application aux corridors

L'échantillon pour les arrêts de la ligne 505 est limité, soit 12 voyages par arrêts, alors que les autres lignes ont un échantillon bien plus grand. N'ayant qu'un seul autobus instrumenté pour le tracé 3 de la ligne 505, la figure 4-31 montre des boîtes à moustaches écrasées ou aplaties et présente donc ce qui s'est passé avec le seul voyage qui a emprunté ce tracé. De plus, à cause du nombre peu élevé de voyages pour le tracé 2 de la ligne 505, soit trois voyages, les boîtes à moustaches à la figure 4-30 sont à prendre avec prudence, les extrémités des moustaches et la médiane étant les seules à avoir un sens.

Sur cette même figure, il est possible de remarquer que les arrêts 12 et 13 sont presque superposés (d'Amos / Saint-Julien et Pie-IX / Arthur-Champoux respectivement). La figure 4-34 montre une partie de la trace GPS du véhicule 26076 qui a été utilisé pour un des voyages du tracé de la ligne 505. La trace montre que le véhicule n'est pas passé par l'arrêt 12 et que le système a tout de même détecté un passage de l'autobus à cet arrêt qui est à environ 200 m du boulevard Pie-IX.

L'arrêt 2 de la ligne 467 en direction sud est dans un cas similaire (voir figure 4-25). La trace GPS du véhicule 29825, dont une partie est présentée à la figure 4-35, révèle que les usagers montent sur le boulevard Saint-Michel juste en amont du boulevard Henri-Bourassa, alors que l'autobus est orienté vers le nord. L'autobus effectue une boucle sur plusieurs rues avant de se diriger vers le sud. Ces cas exposent bien le fait que le tableau de bord permet de faire ressortir les anomalies dans la détection des différents arrêts.

Dans la plupart des lignes-tracés, il est possible d'observer que les boîtes à moustaches de ponctualité forment une espèce de tunnel en V ou un faisceau divergent : au premier arrêt, la boîte est relativement petite et entre dans les normes de ponctualité — 60 secondes d'avance à

180 secondes de retard —, et d'arrêt en arrêt, la boîte s'élargit de plus en plus. L'évolution de la ponctualité des autobus au fil du parcours montré par les tableaux de bord tend à confirmer ce qui a été remarqué dans la section 4.4.

Le planificateur cherche à avoir des autobus qui respectent l'horaire le plus possible, et veut donc que ces boîtes à moustaches restent dans la bande formée par les normes de ponctualité. Le tableau de bord montre dans quelle mesure les autobus sont ponctuels : une boîte à moustaches qui sort de la bande indique un problème à l'arrêt. Par exemple, pour la ligne 67 en direction nord (figure 4-27), plus de la moitié des autobus sont en avance à l'arrêt 18 (Saint-Michel / Jean-Talon), juste avant la station de métro Saint-Michel, car la médiane est clairement supérieure à 60 secondes d'avance. En observant les arrêts précédents, les autobus auraient tendance à être de plus en plus avance jusqu'à l'arrivée à la station de métro, où les chauffeurs attendent l'heure de départ prévue avant de poursuivre.

Les boîtes à moustaches de la durée des haltes aux arrêts permettent de faire ressortir les arrêts où les autobus ont passé le plus de temps. Pour les corridors exemples, la durée des haltes semble corrélée avec le nombre de montées ou de validations. Plus il y a d'usagers qui montent, plus la halte est longue. Il serait intéressant d'inclure les données sur les feux de circulation pour préciser et discriminer les effets des montées et de la programmation des feux sur la durée des haltes.

Les mouvements d'usagers sont les plus élevés aux stations de métro, comme Pie-IX (boulevards Pie-IX / avenue Pierre-de-Coubertin), Saint-Michel et Joliette. Les différences observées entre les données AFC et APC sont élevées aux stations de métro aussi dans la plupart des lignes-tracés. Toutefois, il est à remarquer que cette différence est aussi systématiquement très élevée au premier arrêt des lignes 67, 467 et 139 et pour la ligne 505, direction nord.

Les profils de charges révèlent le type d'utilisation des différentes lignes. Les lignes du corridor Saint-Michel sont utilisées de manière presque relativement uniforme sur toute leur longueur, sauf aux extrémités de la ligne. Par contre, la charge des lignes du corridor Pie-IX croît du nord vers le sud — ou décroît du sud vers le nord — jusqu'à la station de métro Pie-IX.

Bref, les tableaux de bord, en synthétisant les différents indicateurs, font ressortir en un coup d'œil plusieurs informations et tendances à propos des lignes. Alors que l'emplacement des arrêts dans le tableau et la distribution des distances interarrêts donne une indication sur la détection des haltes aux arrêts, le planificateur peut voir en même temps les données opérationnelles sur le respect de l'horaire et sur la durée des haltes ainsi que les données sur la demande avec le nombre de montées et de descentes et avec la charge des lignes. Ainsi, il peut apporter les corrections et les ajustements nécessaires pour améliorer la performance des lignes ou pour améliorer l'information dont il dispose.



Figure 4-25. Tableau de bord de la ligne 467, direction sud

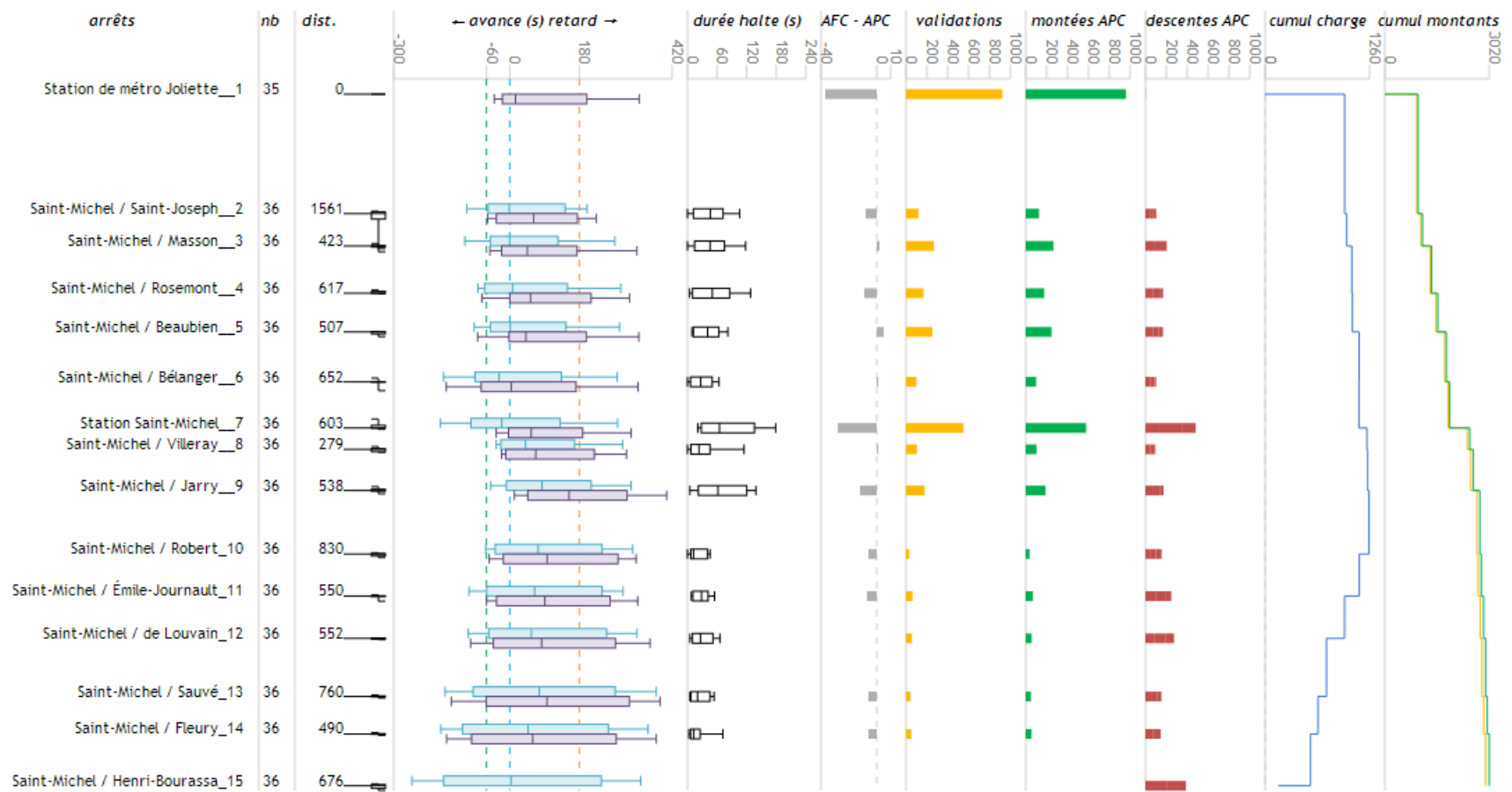


Figure 4-26. Tableau de bord de la ligne 467, direction nord



Figure 4-27. Tableau de bord de la ligne 67, direction nord

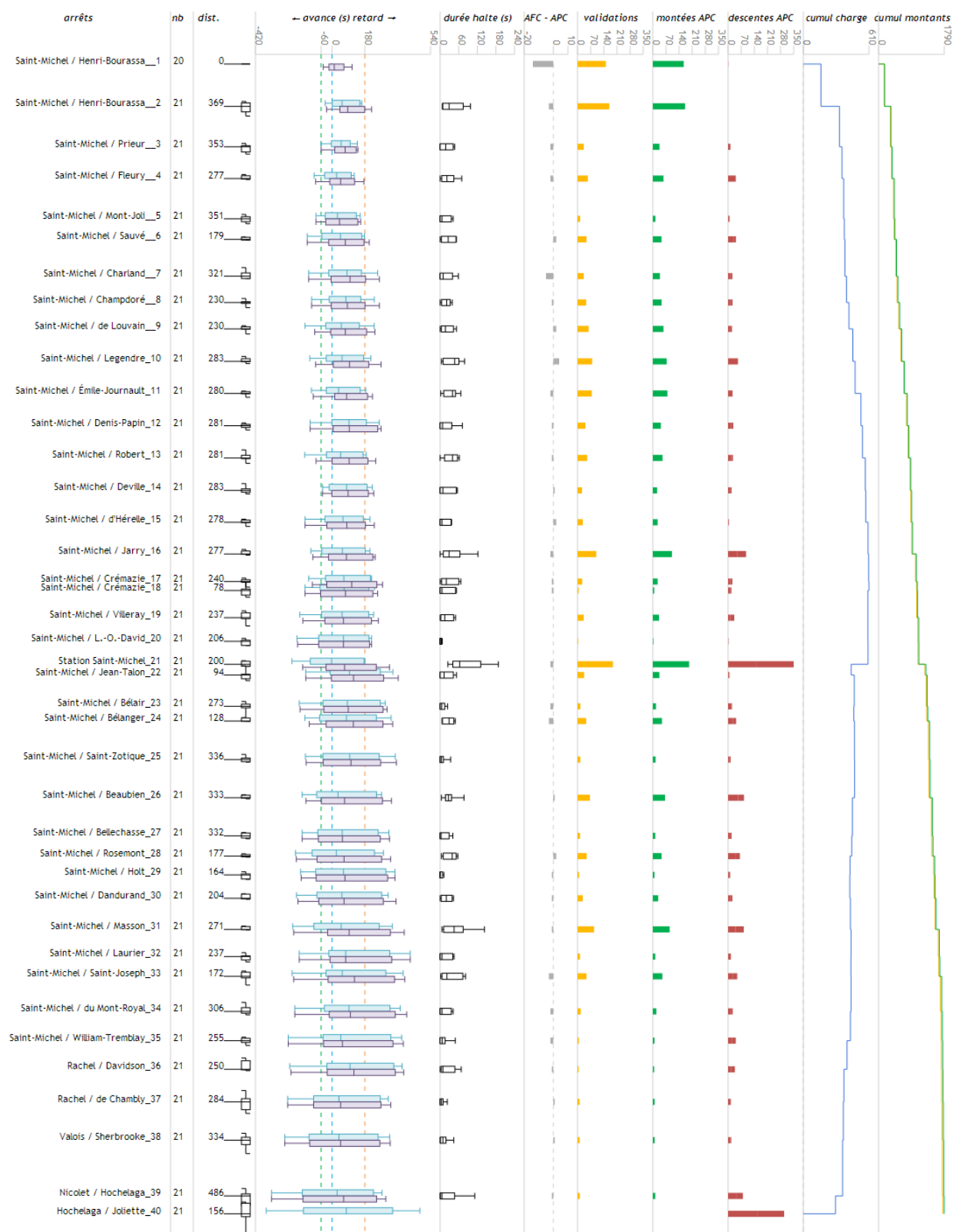


Figure 4-28. Tableau de bord de la ligne 67, direction sud

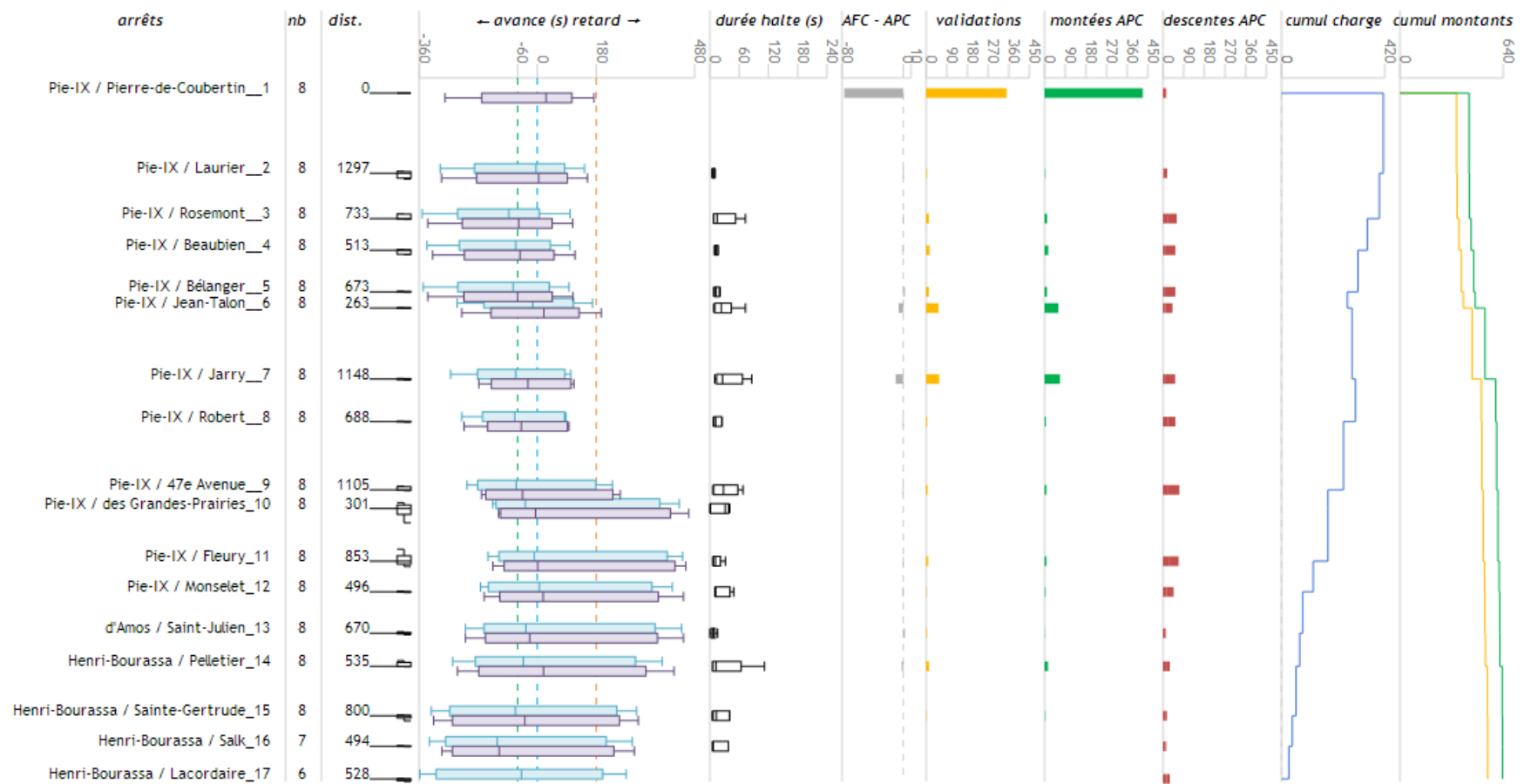


Figure 4-29. Tableau de bord de la ligne 505, direction nord (tracé 1)

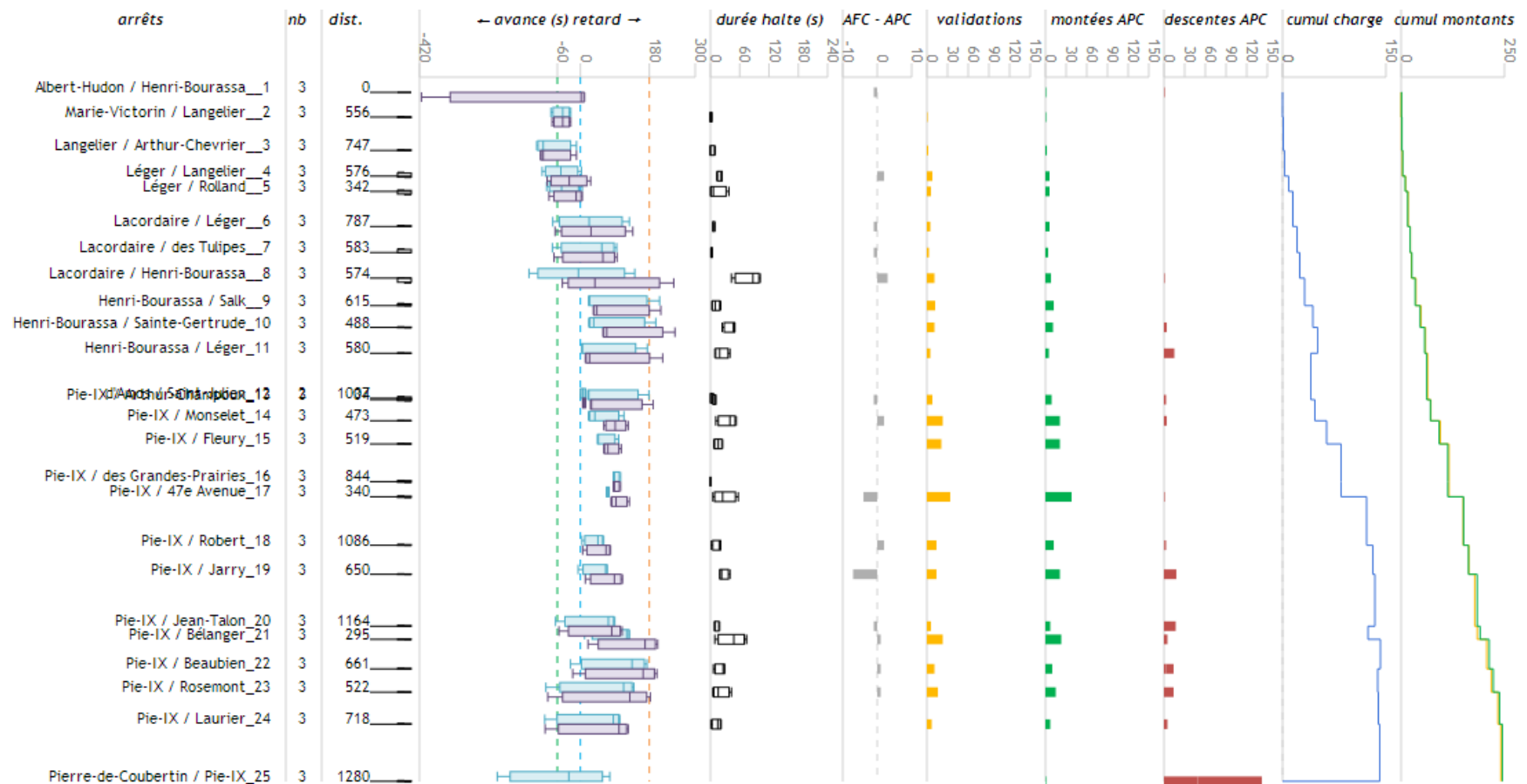


Figure 4-30. Tableau de bord de la ligne 505, direction sud (tracé 2)



Figure 4-31. Tableau de bord de la ligne 505, direction sud (tracé 3)

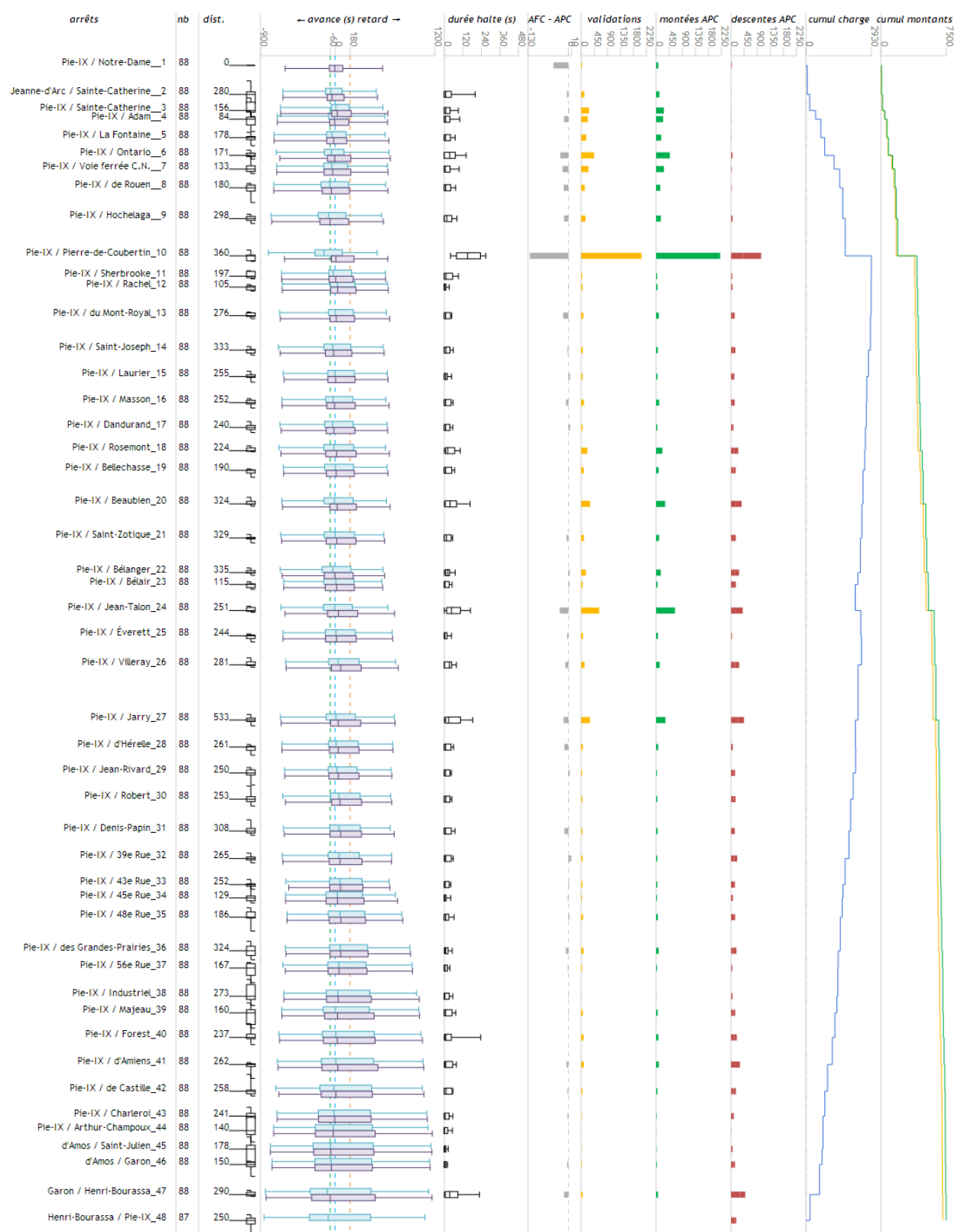


Figure 4-32. Tableau de bord de la ligne 139, direction nord

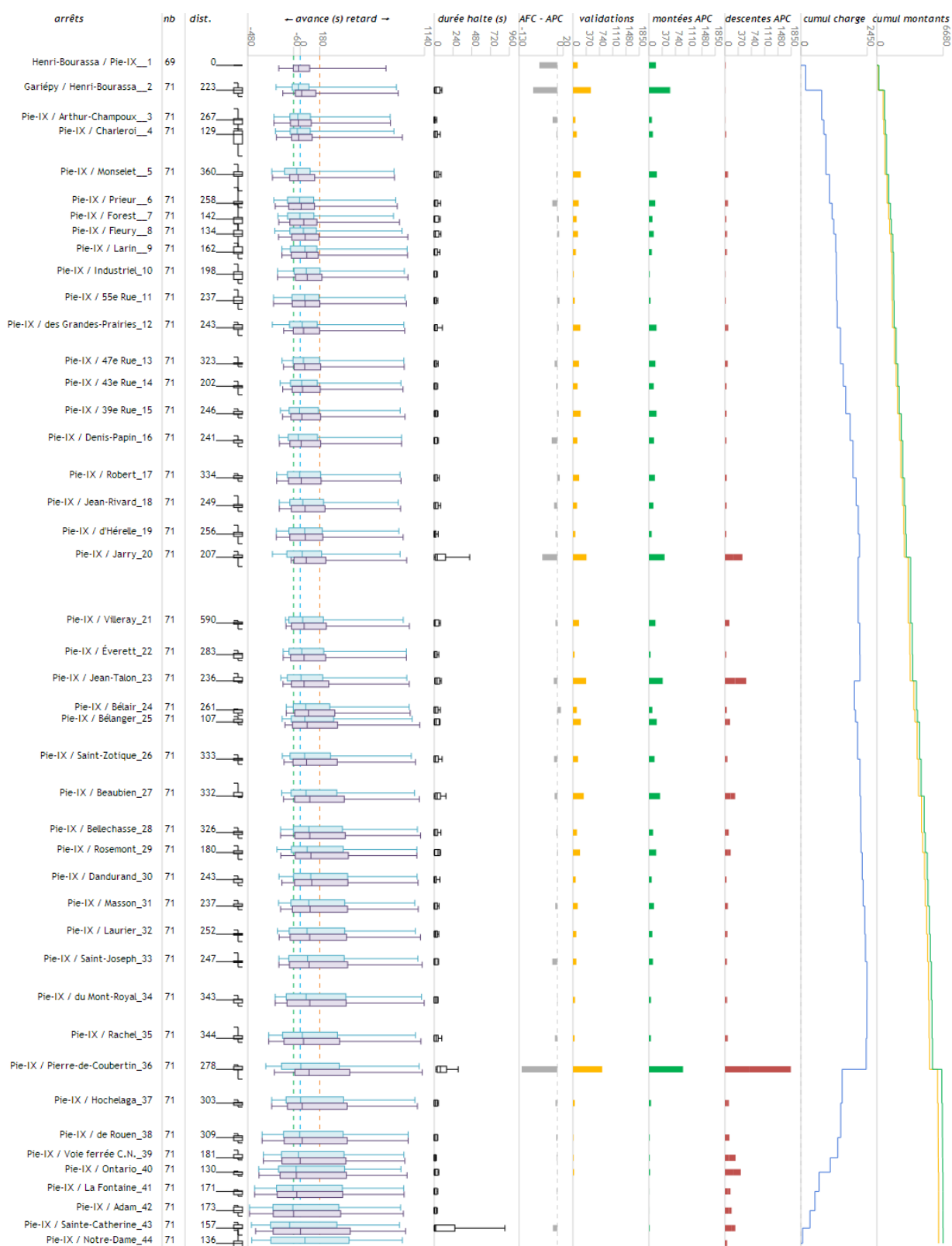
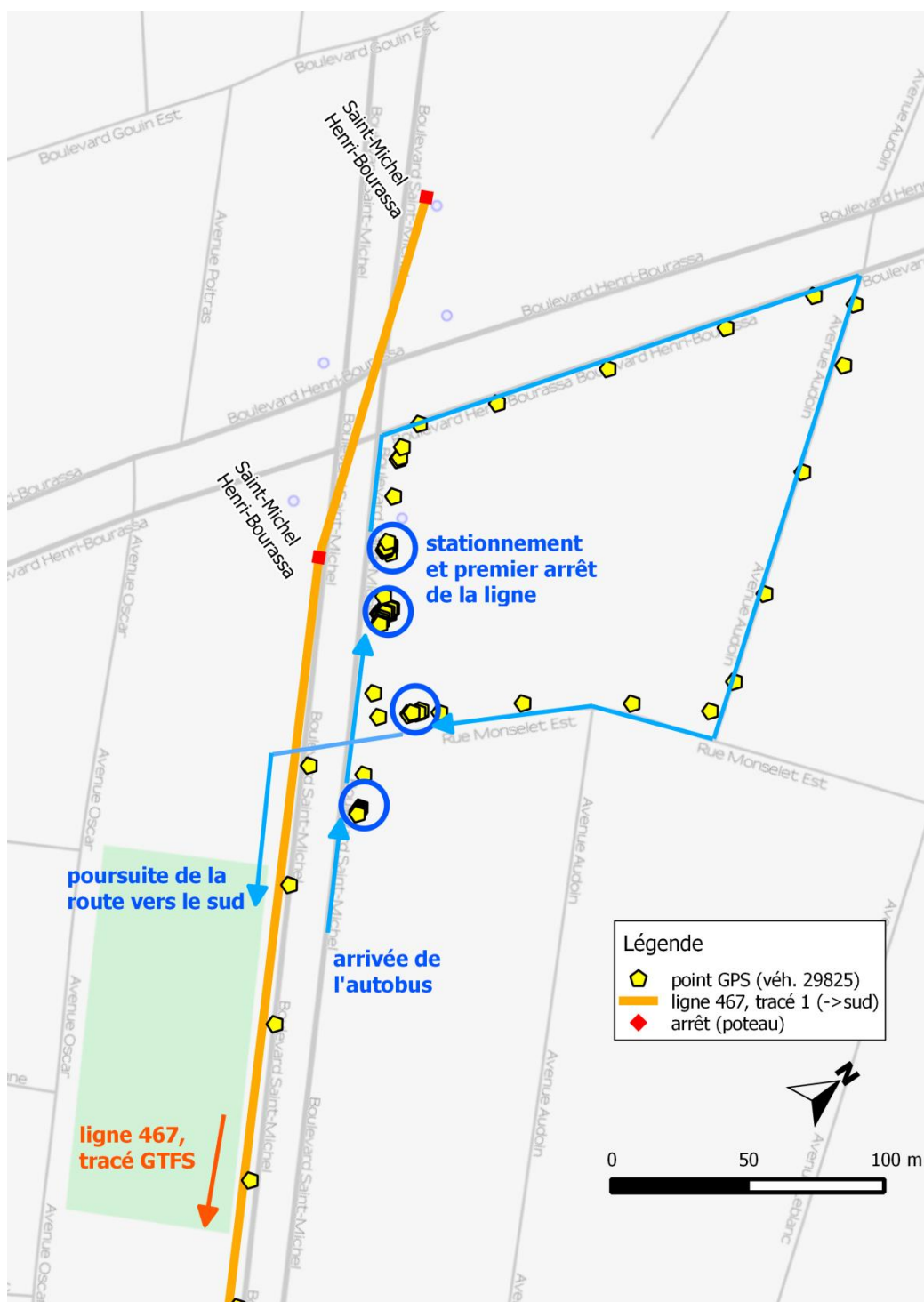


Figure 4-33. Tableau de bord de la ligne 139, direction sud



www.openstreetmap.org/copyright))

Figure 4-34. Trace GPS du véhicule 26076



(fond de carte par © les contributeurs d'OpenStreetMap (www.openstreetmap.org/copyright))

Figure 4-35. Trace GPS du véhicule 29825

4.7 Discussion

D'abord, l'appariement des validations de passages aux arrêts permet la comparaison avec les données des compteurs automatiques, cernant ainsi les arrêts et les moments où les différences entre les deux systèmes sont importantes. Cela peut mener à une réflexion sur la qualité des données et avoir des répercussions importantes sur le contrôle des revenus de l'autorité organisatrice de transport. Les mouvements de passagers aux arrêts peuvent donner l'occasion aux planificateurs d'améliorer l'expérience utilisateur à ces points — abribus, panneaux d'information voyageur en temps réel, etc. Par ailleurs, l'ajout de boîtes à moustaches représentant les mouvements de passagers ou les validations de passages permettrait de mieux décrire l'utilisation des arrêts.

Ensuite, les outils de visualisation des indicateurs pour les corridors d'autobus fournissent aux planificateurs un portrait de l'état du réseau et de son utilisation. L'aspect temporel des outils permet d'évaluer la qualité de l'offre et de prévoir son ajustement — ponctualité, temps de parcours, etc. D'ailleurs, l'ajout d'un filtre par périodes — heures de pointe, midi, soir, etc. — aux tableaux de bord pourrait capturer des phénomènes propres à ces moments de la journée.

En utilisant ces outils avec d'autres corridors d'autobus, les planificateurs pourront les comparer et, par exemple, cibler les axes potentiels d'implantation de lignes à haut niveau de service, comme un service rapide par bus (SRB) ou un système léger sur rail (SLR). De plus, les outils développés peuvent détecter des axes de congestion ou d'incidents. Cette information peut être tenue compte dans la conception des horaires et, en accumulant les données historiques, partagée avec le service de circulation de la ville (Parker, 2008).

Le fait que l'installation du système AVL-APC sur la flotte d'autobus soit encore partielle — environ 16 % des véhicules, et la quantité de voyages instrumentés varient d'une ligne-tracé à l'autre — limite toutefois les informations sur le réseau. En attendant l'installation des équipements sur l'ensemble des autobus vers 2014 dans le cadre du projet iBus (STM, 2012c), Parker (2008) suggère la présélection des voyages qui seront instrumentés si l'utilisation des données doit servir à la planification, ce qui implique une interaction et la coordination avec le service opérationnel de l'autorité de transport. Aussi, d'autres tests sont nécessaires pour valider

les données AVL-APC traitées à l'externe. Il n'est pas exclu de développer une méthodologie pour raffiner le traitement de ces données afin de, par exemple, récupérer des voyages et des arrêts non détectés par le traitement initial.

De plus, les tracés exacts sur le réseau routier du service planifié ne sont pas connus, les lignes-tracés étant déterminées par les données GTFS de poteau à poteau d'arrêts. La disponibilité de telles données ou l'estimation de ces chemins viendraient bonifier l'aspect géomatique d'une fusion de données. Cela aurait pour conséquence, par exemple, une meilleure comparaison entre les distances interarrêts et les endroits où les autobus effectuent leur halte.

Aussi, il faut rappeler le fait que les temps de passage indiqués dans les données GTFS ne sont précis qu'à la minute et qu'il n'y a pas de différence entre les temps de départ et d'arrivée à un arrêt. Le développement d'une méthodologie pour préciser les temps de passage à la seconde — à combiner avec l'évaluation des chemins planifiés sur le réseau — permettrait entre autres de raffiner les estimations sur la ponctualité. En outre, des informations sur les voyages spéciaux planifiés, comme les lignes scolaires, permettraient de compléter le portrait du réseau.

Par contre, les outils développés illustrent tout le potentiel de l'intégration des données des systèmes ADC. Comme le soulignent Boyle (2008) et Parker (2008) (concernant les données AVL et APC), l'utilisation et la fusion des données de plusieurs sources peuvent apporter une meilleure confiance dans les décisions par rapport aux changements de service ou à la révision de la conception de lignes du réseau.

CHAPITRE 5 CONCLUSION

5.1 Résultats

Ce projet consiste à développer des outils et des méthodes systématiques dans la perspective d'exploitation et d'intégration de données de systèmes ADC. Des données des corridors Saint-Michel et Pie-IX ont servi d'exemple d'application des outils créés, en raison de leur achalandage et du nombre de voyages instrumentés.

D'abord, le contexte dans lequel s'inscrit ce mémoire a été exposé. Il a été vu que les méthodes d'enquête traditionnelles, comme les enquêtes OD et les comptages à bord, sont couteuses par rapport aux données qu'elles apportent. Au contraire, les systèmes ADC apportent une quantité importante de données à un coût relativement faible et représentent un grand potentiel d'analyse et de caractérisation de la demande et du réseau. Toutefois, il existe plusieurs défis entourant ces données, notamment sur leur traitement et leur enrichissement. Ensuite, l'organisation générale d'un service d'autobus a été vue : tournée, voyage, ligne-direction-tracé, etc.

Après, les données GTFS, initialement destinées à de l'information voyageur, ont été compressées et une liste des lignes-tracés a été extraite. En estimant les ressources impliquées et son évolution dans le temps, la classification des lignes révèle l'organisation du service et compare l'importance de chaque type de ligne. Ensuite, les données AVL-APC sur 16 % de la flotte d'autobus, traitées par une entreprise externe, permettent d'associer les temps de passage et les mouvements de passagers aux arrêts.

Aussi, les données AFC sont prétraitées en retirant les doublons, les refus et les validations tests. Toutefois, les validations ne sont pas nécessairement associées au bon voyage, car les chauffeurs d'autobus peuvent ne pas changer et ne pas valider le voyage qu'ils s'appêtent à effectuer. Ce problème est contourné en appariant les validations de passages aux arrêts en tenant compte du numéro de véhicule et des temps de passage détectés avec les traces GPS. Pour ce faire, il est supposé qu'un usager valide son passage entre l'arrivée de l'autobus à son arrêt et l'arrivée au prochain arrêt. Les modèles de régression suggèrent que, pour les voyages qui ont des données AFC et APC, les compteurs estiment environ 5 % plus de montées que les validations. Ce phénomène est observable sur l'ensemble du réseau peu importe l'heure du jour. Toutefois, des

différences extrêmes entre ces données sont plus fréquentes en heure de pointe. L'analyse d'un véhicule sur la ligne 139 montre que ces grandes différences sont observées à certains arrêts à certains moments de la journée, en après-midi dans ce cas.

Puis, l'estimation de la ponctualité au début et en fin de voyage montre que les autobus tendent à arriver en avance à la fin d'un HLP (sans tenir compte du temps de recouvrement). Les voyages productifs débutent généralement à l'heure, alors que la distribution de la ponctualité en fin de voyage est étendue avec une légère tendance à arriver en avance.

Ensuite, un diagramme distance-temps montre les voyages effectués dans un corridor donné et les compare avec ceux planifiés. Le côté interactif de l'outil permet à l'utilisateur d'afficher ou de masquer certains voyages, la charge et les mouvements de passagers aux arrêts. Enfin, le tableau de bord permet de visualiser en un coup d'œil plusieurs indicateurs et leurs tendances aux arrêts d'une ligne-tracé : distances interarrêts, ponctualité, durée de halte, mouvements de passagers, charge, etc.

5.2 Perspectives de recherche

Cette étude n'est que quelques pas vers l'intégration de données ADC. D'autres corridors ou axes d'autobus pourraient être examinés pour, entre autres, cibler des axes d'intervention ou d'établissement de service à fréquence élevée. Bien sûr, l'examen de l'ensemble du réseau de la STM ne sera véritablement possible que lorsque l'implantation du SAEIV — projet iBus — sera généralisée.

Aussi, en plus d'améliorer la dimension géomatique avec les tracés exacts sur le réseau routier, les données GTFS pourraient bénéficier d'un enrichissement sur le plan temporel. Cela permettrait une meilleure comparaison avec les temps de passage effectivement réalisés. De plus, la dimension géomatique pourrait être améliorée avec une estimation des chemins sur le réseau routier pour les voyages planifiés dans le but de mieux comparer les distances interarrêts mesurées. L'inclusion des lignes spéciales, pour les écoliers et les industries entre autres, permettrait de compléter le réseau planifié.

Par ailleurs, l'optimisation des haut-le-pied est un enjeu de taille étant donné qu'ils sont improductifs — ne transportant normalement aucun passager, ces voyages ne génèrent aucun revenu en principe — et qu'ils pourraient représenter en gros entre 15 % et 20 % des véhicules-

heures (voir section 3.1.3). Ces données pourraient servir à analyser ces voyages et à vérifier l'optimisation des tournées.

Enfin, la confrontation entre les données APC et AFC soulèvent des questions notamment sur le contrôle des revenus. Cela constitue une préoccupation importante pour une autorité organisatrice de transport, alors que la clientèle forme en 2012 près de 46 % des revenus de la Société de transport de Montréal (2013b). Une analyse plus poussée de ces données apporterait un éclairage plus détaillé de la situation.

De plus, d'autres sources de données pourraient être ajoutées et incluses. Par exemple, les données provenant de systèmes ADC ne tiennent souvent pas compte de l'aspect socioéconomique des déplacements, comme le fait remarquer Bagchi et White (2005) pour certaines données de cartes à puce. L'intégration d'autres sources de données comme les enquêtes OD ou les recensements pourrait ajouter cette dimension en apportant certains attributs sur les usagers, comme l'âge, le revenu ou la possession automobile. Il serait aussi intéressant d'inclure des données sur les feux de circulation. En les incluant avec un diagramme distance-temps et en les comparant aux temps de passage, ces données pourraient contribuer à expliquer les temps passés aux arrêts.

Avec un jeu de données AFC-AVL de la Société de transport de l'Outaouais (STO), Chu (2010) montre qu'une grande base de données permet de reconstruire les itinéraires des déplacements des usagers en utilisant une approche multijours. Cette méthode pourrait être envisagée à plus grande échelle sur un plus grand réseau comme celui de la STM. En outre, l'analyse de données étalées sur plusieurs jours permettrait d'identifier des phénomènes propres à certains moments de l'année (effets de saisonnalité, congés, etc.).

Le haut niveau de résolution des données ADC laisse aussi la possibilité de créer une microsimulation du réseau d'autobus. La microsimulation automobile est largement répandue. Par exemple, Spurr (2010) modifie la plateforme TRANSIMS (Transportation Analysis Simulation System) pour évaluer l'équité interrégionale de l'utilisation des ponts reliant Montréal à ses banlieues. Par contre, la microsimulation d'un réseau multimodal est encore au stade embryonnaire (Volosin, Paul, Pendyala, Grady, & Gardner, 2012). L'inclusion d'un réseau d'autobus dans une microsimulation comporte certains défis, comme la prise en compte de

l'interaction avec les autres véhicules, mais pourrait dessiner un portrait représentatif de ce qui se passe sur le réseau.

Néanmoins, les outils développés dans ce projet montrent le potentiel de l'intégration de données de sources multiples. D'une part, ils permettent de bonifier la compréhension de l'offre autant annoncée que réalisée. D'autre part, la demande et son interaction avec le réseau peut être finement examinée.

Ces informations offre une opportunité aux autorités de transport de raffiner les méthodes de planification, ce qui devrait améliorer la performance et l'efficacité du réseau. Ces enjeux s'expliquent très bien d'un point de vue financier : même de simples petites modifications sur une ligne, opérée de manière récurrente à longueur d'année, peut avoir un impact très important sur les dépenses de l'AOT.

De plus, un réseau efficace est généralement connu pour satisfaire davantage les usagers. Si cela se traduit par une hausse de l'achalandage, le transport en commun peut ainsi mieux remplir son rôle d'élément clé dans la poursuite du développement durable, dans le contexte des changements climatiques.

BIBLIOGRAPHIE

- Acumen Building Enterprise Inc., & Booz Allen Hamilton Inc. (2006). *Smartcard interoperability issues for the transit industry* (Rapport n° R-115). Transit Cooperative Research Program. Tiré de <http://pubsindex.trb.org/view/2006/M/796099>
- Agence métropolitaine de transport de Montréal. (2010). *Enquête Origine-Destination 2008 — La mobilité des personnes dans la région de Montréal : faits saillants* (Rapport n° 2008). Tiré de http://www.enquete-od.qc.ca/docs/EnqOD08_FaitsSaillants.pdf
- Atkins, S. T. (1986). Transportation planning models —what the papers say. *Traffic Engineering and Control*, 27(9), 460-467.
- Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, 12(5), 464-474. doi: 10.1016/j.tranpol.2005.06.008
- Bates, J. (2008). History of demand modelling. Dans D. A. Hensher & K. J. Button (Édit.), *Handbook of transport modelling* (2^e éd., p. 11-34). Amsterdam, Pays-Bas: Elsevier.
- Beimborn, E. A., Kennedy, R., & Schaefer, W. (1996). *Inside the blackbox: making transportation models work for livable communities*. Milwaukee, WI: Citizens for a Better Environment.
- Bierlaire, M. (1998). Discrete choice models. Dans M. Labbé, G. Laporte, K. Tanczos & P. Toint (Édit.), *Operations research and decision aid methodologies in traffic and transportation management* (vol. 166, p. 203-227). Berlin, Allemagne: Springer.
- Bostock, M. (2012). Data-Driven Documents (d3.js) [Bibliothèque Javascript]. Tiré de <http://d3js.org/>
- Bourdon, R., & Leclerc, H. (2013). WampServer (Version 2.4) [Environnement WAMP]. Tiré de <http://www.wampserver.com/>
- Box plot. (s.d.). Dans *Wikipedia*. Consulté le 11 décembre 2013, tiré de http://en.wikipedia.org/wiki/Box_plot
- Boyle, D. K. (2008). *Passenger Counting Systems* (Rapport n° S-77). Transit Cooperative Research Program. Tiré de http://www.tcrponline.org/bin/publications.pl?mode=abstract&cat_id=23&pub_id=1436
- Brumfield, G. (2011). High-energy physics: down the petabyte highway. *Nature*, 469(7330), 282-283. doi: 10.1038/469282a

- Bus bunching. (s.d.). Dans *Wikipedia*. Consulté le 22 juillet 2013, tiré de http://en.wikipedia.org/wiki/Bus_bunching
- Chapleau, R. (1986). *Transit network analysis and evaluation with a totally disaggregate approach*. Communication présentée à World Conference on Transportation Research, Vancouver, BC.
- Chapleau, R. (1992). *La modélisation de la demande de transport urbain avec une approche totalement désagrégée*. Communication présentée à World Conference on Transportation Research, Lyon, France.
- Chapleau, R., & Allard, B. (2010). *Merging AFC, APC, GPS and GIS-T data to generate productivity indicators and travel demand models in public transit*. Communication présentée à 12th World Conference on Transportation Research, Lisbonne, Portugal.
- Chapleau, R., Allard, B., & Sarrazin-Canova, M. (1982). *MADITUC, un modèle de planification opérationnelle adapté aux entreprises de transport en commun de taille moyenne* (Rapport n° 265). Montréal, QC: Centre de recherche sur les transports de l'Université de Montréal.
- Chapleau, R., Chu, K. K., & Allard, B. (2011). *Synthesizing AFC, APC, GPS and GIS data to generate performance and travel demand indicators for public transit*. Communication présentée à Transportation Research Board 90th Annual Meeting, Washington, DC. Tiré de <http://pubsindex.trb.org/view/2011/C/1092733>
- Cheung, F. (2006). Implementation of nationwide public transport smart card in the Netherlands: cost-benefit analysis. *Transportation Research Record: Journal of the Transportation Research Board*(1971), 127-132. doi: 10.3141/1971-17
- Chu, K. K. (2010). *Leveraging data from a smart card automatic fare collection system for public transit planning*. (Thèse de doctorat, École Polytechnique de Montréal, Montréal, QC). Accessible par PolyPublie. (441). Tiré de <http://publications.polymtl.ca/441/>
- Chu, K. K., & Chapleau, R. (2007). *Imputation techniques for missing fields and implausible values in public transit smart card data*. Communication présentée à 11th World Conference on Transportation Research, Berkeley, CA.
- Chu, K. K., & Chapleau, R. (2008). Enriching archived smart card transaction data for transit demand modeling. *Transportation Research Record: Journal of the Transportation Research Board*(2063), 63-72. doi: 10.3141/2063-08
- Chu, K. K., & Chapleau, R. (2010). Augmenting transit trip characterization and travel behavior comprehension: multiday location-stamped smart card transactions. *Transportation*

Research Record: Journal of the Transportation Research Board(2183), 29-40. doi: 10.3141/2183-04

Chu, K. K., Chapleau, R., & Trépanier, M. (2009). Driver-assisted bus interview: passive transit travel survey with smart card automatic fare collection system and applications. *Transportation Research Record: Journal of the Transportation Research Board*(2105), 1-10. doi: 10.3141/2105-01

Desharnais, M.-C. (2009). *Caractérisation objective de la demande de transport adapté*. (Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, QC). Accessible par ProQuest Dissertations & Theses. Tiré de <http://search.proquest.com/pqdtft/docview/305138444/abstract/13429C9BBFA4A4AC439/16?accountid=40695>

Diab, E. I., & El-Geneidy, A. M. (2012). Understanding impacts of a combination of service improvement strategies on bus running time and passengers' perception. *Transportation Research Part A: Policy and Practice*, 46(3), 614-625. Tiré de <http://trid.trb.org/view.aspx?id=1134361>

Domencich, T. A., & McFadden, D. (1975). *Urban travel demand: a behavioral analysis*. Amsterdam, Pays-Bas; New York, NY: North-Holland Publishing Company; American Elsevier Publishing Company.

Doucet-Koussaya, W. (2012). *Analyse exploratoire de transactions de cartes à puces [sic] sur le réseau du métro de Montréal*. (Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, QC). Accessible par PolyPublie. (875). Tiré de <http://publications.polymtl.ca/875/>

Douglass, M., & Dryden, J. (2012). *Transportation corridors and community structures* (Rapport n° 496). New Zealand Transport Agency. Tiré de <http://nzta.govt.nz/resources/research/reports/496/index.html>

Dumbill, E. (2012). What is big data? [Billet de blogue]. Tiré de <http://strata.oreilly.com/2012/01/what-is-big-data.html>

El-Geneidy, A. M., & Surprenant-Legault, J. (2010). *Limited stop bus service: an evaluation of an implementation strategy*. Communication présentée à Transportation Research Board 89th Annual Meeting, Washington, DC.

Eom, J. K., Choi, M. H., & Lee, J. (2012). *Evaluation of metro service quality using transit smart card data*. Communication présentée à Transportation Research Board 91st Annual Meeting, Washington, DC. Tiré de <http://amonline.trb.org/1se9fo/1>

- Faivre d'Arcier, B. (2012). *Mesure de la performance des lignes de transport public urbain — Projet APEROL — Amélioration de la performance économique des réseaux par l'optimisation des lignes* (Rapport n° 20371). Laboratoire d'économie des transports & Groupe opérationnel 6 du Programme de recherche et d'innovation dans les transports terrestres. Tiré de <http://temis.documentation.developpement-durable.gouv.fr/document.xsp?id=Temis-0076754>
- Furth, P. G., Hemily, B., Muller, T. H. J., & Strathman, J. G. (2006). *Using archived AVL-APC data to improve transit performance and management* (Rapport n° R-113). Transit Cooperative Resesarch Program. Tiré de http://www.tcrponline.org/bin/publications.pl?mode=abstract&cat_id=23&pub_id=1267
- Gerson, A. (2011). *Analyse comparée de corridors importants de transport de surface à Montréal*. (Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, QC). Accessible par PolyPublie. (512). Tiré de <http://publications.polymtl.ca/512/>
- Goodwin, P. B. (1984). *Simulation as a tool for investigating model bias* (Rapport n° TSU--233). Oxford, Royaume-Uni: University of Oxford.
- Google. (2012a). What is GTFS? *Google Developers — Transit*. Tiré de <https://developers.google.com/transit/gtfs/>
- Google. (2012b). General Transit Feed Specification. *Google Developers — Transit*. Tiré de http://code.google.com/intl/en/transit/spec/transit_feed_specification.html
- Graham, M. (2012). Big data and the end of theory? [Billet de blogue]. Tiré de <http://www.guardian.co.uk/news/datablog/2012/mar/09/big-data-theory>
- Hensher, D. A., & Button, K. J. (2007). *Handbook of Transport Modelling* (2^e éd.). Bingley, Royaume-Uni: Emerald Group Publishing Limited.
- Hines, W. W., Montgomery, D., Goldsman, D. M., & Borror, C. M. (2005). *Probabilités et statistique pour ingénieurs* (L.-D. Adjengue, J.-P. Carmichael, L. Collet & J. L. Massé, Trad., 1^e éd.). Montréal, QC: Chenelière Éducation.
- Init Innovation in traffic systems AG. (2011). Projects — North America: Montréal, Canada. Tiré de <http://www.initag.com/share/projects/en/Montreal.pdf>
- Iseki, H., Yoh, A. C., & Taylor, B. D. (2007). Are smart cards the smart way to go? Examining their adoption by U.S. transit agencies. *Transportation Research Record: Journal of the Transportation Research Board*(1992), 45-53. doi: 10.3141/1992-06

- Jang, W. (2010). Travel time and transfer analysis using transit smart card data. *Transportation Research Record: Journal of the Transportation Research Board*(2144), 142-149. doi: 10.3141/2144-16
- Jones, P. M., Dix, M. C., Clarke, M. I., & Heggie, I. G. (1983). Understanding travel behaviour. *Journal of Forecasting*, 4(3), 315-316.
- Kikuchi, S., Mangalpally, S., & Gupta, A. (2006). Method for balancing observed boarding and alighting counts on a transit line. *Transportation Research Record: Journal of the Transportation Research Board*(1971), 42-50. doi: 10.3141/1971-07
- Lomone, A., & Chapleau, R. (2013). *Convergence de données GTFS, GPS, de compteurs de passagers et de cartes à puce et application à des corridors de bus*. Communication présentée à 48e congrès de l'Association québécoise des transports, Montréal, QC.
- Manheim, M. L. (1979). *Fundamentals of transportation systems analysis* (vol. 1). Cambridge, MA: MIT Press.
- McNally, M. (2008). The four-step model. Dans D. A. Hinsher & K. J. Button (Édit.), *Handbook of transport modelling* (2^e éd., p. 35-53). Amsterdam, Pays-Bas: Elsevier.
- Ministère des Transports du Québec. (2009). Plan stratégique 2008 - 2012. Tiré de http://www.mtq.gouv.qc.ca/portal/page/portal/Librairie/Publications/fr/ministere/plan_strat_2008_2012.pdf
- Ministère des Transports du Québec. (2012). Amélioration des services en transport en commun. Tiré de http://www.mtq.gouv.qc.ca/portal/page/portal/ministere/ministere/programmes_aide/transport_collectif/amel_serv_trans_com
- Ministère du Développement durable, de l'Environnement et des Parcs du Québec. (2006). Plan d'action 2006-2012 sur les changements climatiques. Tiré de http://www.mddep.gouv.qc.ca/changements/plan_action/index-mesures.htm
- Ministère du Développement durable, de l'Environnement et des Parcs du Québec. (2012). Le Québec en action — Plan d'action 2013-2020 sur les changements climatiques. Tiré de http://www.mddefp.gouv.qc.ca/changements/plan_action/index.htm
- Ministère du Développement durable, de l'Environnement et des Parcs du Québec,. (2013). Inventaire québécois des émissions de gaz à effet de serre en 2010 et leur évolution depuis 1990. Tiré de <http://www.mddep.gouv.qc.ca/changements/ges/>

- Morency, C. (2004). *Contributions à la modélisation totalement désagrégée des interactions entre mobilité urbaine et dynamiques spatiales*. (Thèse de doctorat, École Polytechnique de Montréal, Montréal, QC). Accessible par ProQuest Dissertations & Theses. Tiré de <http://search.proquest.com/pqdtft/docview/305062596/abstract/1342EDAE8006941AB1A/1?accountid=40695>
- O'Donnell, E., & Smith David, J. (2000). How information systems influence user decisions: a research framework and literature review. *International Journal of Accounting Information Systems*, 1(3), 178-203. doi: 10.1016/S1467-0895(00)00009-9
- Office québécois de la langue française. (2002). Embarquer et débarquer. Dans *Banque de dépannage linguistique*. Tiré de http://bdl.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?t1=1&id=3806
- Office québécois de la langue française. (2008). Orienté objet. Dans *Le grand dictionnaire terminologique*. Tiré de http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8358467
- Ortúzar, J. d. D., & Willumsem, L. G. (2011). *Modelling transport* (4^e éd.). Oxford: Wiley-Blackwell.
- Park, J. Y., Kim, D.-J., & Lim, Y. (2008). Use of smart card data to define public transit use in Seoul, South Korea. *Transportation Research Record: Journal of the Transportation Research Board*(2063), 3-9. doi: 10.3141/2063-01
- Parker, D. J. (2008). *AVL Systems for bus transit: update* (Rapport n° S-73). Transit Cooperative Research Program. Tiré de http://www.tcrponline.org/bin/publications.pl?mode=abstract&cat_id=23&pub_id=1400
- Pelletier, M.-P., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: a literature review. *Transportation Research Part C: Emerging Technologies*, 19(4), 557-568. doi: 10.1016/j.trc.2010.12.003
- Public Utility Research Center. (2006). Terminology and Basic Relationships. Tiré de http://warrington.ufl.edu/centers/purc/docs/resources_Terminology_PublicTransport.pdf
- Roth, M. (2010). How Google and Portland's TriMet set the standard for open transit data [Billet de blogue]. Tiré de <http://sf.streetsblog.org/2010/01/05/how-google-and-portlands-trimet-set-the-standard-for-open-transit-data/>
- Seaborn, C., Attanucci, J., & Wilson, N. H. M. (2009). Analyzing multimodal public transport journeys in London with smart card fare payment data. *Transportation Research Record: Journal of the Transportation Research Board*(2121), 55-62. doi: 10.3141/2121-06

Seven-number summary. (s.d.). Dans *Wikipedia*. Consulté le 26 juin 2013, tiré de http://en.wikipedia.org/wiki/Seven-number_summary

Slavin, H., Rabinowicz, A., Brandon, J., Flammia, G., & Freimer, R. (2009). Using automatic fare collection data, GIS, and dynamic schedule queries to improve transit data and transit assignment models. Dans Nigel H.M. Wilson and Agostino Nuzzolo (Édit.), *Schedule-Based Modeling of Transportation Networks* (p. 101-118). New York, NY: Springer.

Smith, S. A. (1999). *Guidebook for transportation corridor studies: a process for effective decision-making* (Rapport n° R-435). Washington, DC: National Cooperative Highway Research Program.

Société de transport de Montréal. (2008). Lancement de la carte OPUS dans les transports collectifs au Québec. Tiré de <http://www.stm.info/fr/presse/communiques/2008/lancement-de-la-carte-opus-----dans-les-transports-collectifs-au-quebec>

Société de transport de Montréal. (2010a). Rapport d'activités 2009. Tiré de <http://www.stm.info/sites/default/files/ra2009.pdf>

Société de transport de Montréal. (2010b). Programme triennal d'immobilisations 2011 — 2013. Tiré de http://www.stm.info/sites/default/files/pti_11-13.pdf

Société de transport de Montréal. (2011a). Rapport d'activités 2010. Tiré de <http://www.stm.info/sites/default/files/ra2010.pdf>

Société de transport de Montréal. (2011b). Programme triennal d'immobilisations 2012 — 2014. Tiré de http://www.stm.info/sites/default/files/pti_12-14.pdf

Société de transport de Montréal. (2012a). Chronologie de l'entreprise. Tiré de <http://www.stm.info/fr/a-propos/decouvrez-la-STM-et-son-histoire/histoire/chronologie-de-lentreprise>

Société de transport de Montréal. (2012b). Rapport d'activités 2011. Tiré de <http://www.stm.info/sites/default/files/ra2011.pdf>

Société de transport de Montréal. (2012c). La STM acquiert un système d'information en temps réel pour son réseau de surface. Tiré de <http://www.stm.info/fr/presse/communiques/2012/la-stm-acquiert-un-systeme-d-information-en-temps-reel-pour-son-reseau-de-surface>

Société de transport de Montréal. (2012d). Programme triennal d'immobilisations 2013 — 2015. Tiré de http://www.stm.info/sites/default/files/pdf/fr/pti_13-15.pdf

- Société de transport de Montréal. (2013a). R-105 : Règlement concernant les conditions au regard de la possession et de l'utilisation de tout titre de transport émis par la Société de transport de Montréal. *Règlements*. Tiré de <http://www.stm.info/fr/a-propos/informations-entreprise-et-financieres/les-reglements-et-politiques/reglement-r-105>
- Société de transport de Montréal. (2013b). Rapport d'activités 2012. Tiré de <http://www.stm.info/sites/default/files/pdf/fr/ra2012.pdf>
- Société de transport de Montréal. (2013c). Famille de services — À chacun son bus. Tiré de <http://www.stm.info/fr/infos/reseaux/bus>
- Spurr, T. (2010). *The assignment of trips to a road network for the analysis of equitable transport*. (Thèse de doctorat, École Polytechnique de Montréal, Montréal, QC). Accessible par PolyPublie. (473). Tiré de <http://publications.polymtl.ca/473/>
- Stopher, P. R. (1983). *Data needs and data collection — state of the practice*. Communication présentée à Travel analysis methods for the 1980's, Easton, MD.
- Strathman, J. G., Kimpel, T. J., Broach, J., Wachana, P., Coffel, K., Callas, S., . . . Elmore-Yalch, R. (2008). *Leveraging ITS Data for transit market research: a practitioner's guidebook* (Rapport n° R-126). Transit Cooperative Resesarch Program. Tiré de http://www.tcrponline.org/bin/publications.pl?mode=abstract&cat_id=23&pub_id=1402
- Surprenant-Legault, J., & El-Geneidy, A. M. (2011). Introduction of a reserved bus lane: impact on bus running time and on-time performance. *Transportation Research Record: Journal of the Transportation Research Board*(2218), 10-18.
- Tétreault, P. R. (2009). *Designing a parallel limited stop service bus route in montréal using automatic vehicule location and passenger counting data*. (Mémoire de maitrise, McGill University, Montréal, QC). Accessible par eScholarship@McGill. Tiré de http://digitool.Library.McGill.CA:80/R/-?func=dbin-jump-full&object_id=92918&silo_library=GEN01
- Tétreault, P. R., & El-Geneidy, A. M. (2009). *Estimating bus run times for new limited-stop service using archived AVL and APC data*. Communication présentée à Transportation Research Board 88th Annual Meeting, Washington, DC. Tiré de http://tram.mcgill.ca/Research/Publications/STM_67_express.pdf
- The Center for Transit-Oriented Development. (2010). *Transit corridors and TOD: connecting the dots is important* (Rapport n° TOD 203). The Center for Transit-Oriented Development. Tiré de <http://www.reconnectingamerica.org/resource-center/browse-research/2010/tod-203-transit-corridors-and-tod/>

- The Economist. (25 février 2010a). Data, data everywhere. *The Economist*, 394(8671), S.3-S.5. Tiré de <http://www.economist.com/node/15557443>
- The Economist. (25 février 2010b). The data deluge. *The Economist*, 394(8671), 11. Tiré de <http://www.economist.com/node/15579717>
- Thévenin, T. (2001). *L'analyse désagrégée : un luxe ou une nécessité?* Communication présentée à 5es rencontres de Théo Quant, Besançon, France. Tiré de <http://thema.univ-fcomte.fr/theoq/pdf/2001/TQ2001%20ARTICLE%2023.pdf>
- Thomas, I. (2001). Cartographie d'aujourd'hui et de demain : rappels et perspectives. *Cybergeo : Revue européenne de géographie*(189), 1-23. doi: 10.4000/cybergeo.3812
- Transportation Research Board, & Kittelson and Associates Incorporated. (2003). *Transit capacity and quality of service manual* (2^e éd.). Washington, D.C.: Transportation Research Board.
- Treiber, M., & Kesting, A. (2013). *Traffic Flow Dynamics*. Berlin, Allemagne: Springer.
- Trépanier, M. (1999). *Modélisation totalement désagrégée et orientée-objet appliquée aux transports urbains*. (Thèse de doctorat, École Polytechnique de Montréal, Montréal, QC). Accessible par ProQuest Dissertations & Theses. Tiré de <http://search.proquest.com/pqdtft/docview/304574670/abstract/13692AA26FD67F76F92/1?accountid=40695>
- Trépanier, M. (2011). L'exploitation des données de cartes à puce à des fins de planification des transports collectifs urbains. *Recherche Transports Sécurité*, 27. doi: 10.1007/s13547-011-0019-z
- Trépanier, M., Barj, S., Dufour, C., & Poilpré, R. (2004). *Examen des potentialités d'analyse des données d'un système de paiement par carte à puce en transport urbain*. Communication présentée à Congrès annuel 2004 de l'Association des transports du Canada, Québec, QC. Tiré de <http://www.tac-atc.ca/english/resourcecentre/readingroom/conference/conf2004/docs/s8/dufour.pdf>
- Trépanier, M., & Chapleau, R. (1996). *Un modèle d'analyse désagrégée des générateurs de déplacements : vers une connaissance détaillée de l'utilisation des lieux urbains*. Communication présentée à 31e congrès de l'Association québécoise du transport et des routes, Québec, QC.
- Trépanier, M., & Chapleau, R. (2001). Analyse orientée-objet et totalement désagrégée des données d'enquêtes ménages origine-destination. *Revue canadienne de génie civil*, 28(1), 48-58. Tiré de <http://www.nrcresearchpress.com/doi/abs/10.1139/100-106>

- Trépanier, M., Chapleau, R., Allard, B., & Morency, C. (2004). *Information technologies for household survey management*. Communication présentée à Annual Conference of the Transportation Association of Canada, Québec, QC. Tiré de <http://www.tac-atc.ca/english/resourcecentre/readingroom/conference/conf2004/docs/s6/trepanier.pdf>
- Trépanier, M., Chapleau, R., & Morency, C. (2008). Tools and methods for a transportation household survey. *Urban and Regional Information Systems Association Journal*, 20(1), 35-43. Tiré de <http://urisa.org/files/Trepanier%20Chapleau%20Morency%20Vol%2020%20No%201.pdf>
- Trépanier, M., Morency, C., & Agard, B. (2009). Calculation of transit performance measures using smartcard data. *Journal of Public Transportation*, 12(1), 79-96.
- Uniman, D. L., Attanucci, J., Mishalani, R. G., & Wilson, N. H. M. (2010). Service reliability measurement using automated fare card data: application to the London Underground. *Transportation Research Record: Journal of the Transportation Research Board*(2143), 92-99. doi: 10.3141/2143-12
- Utsunomiya, M., Attanucci, J., & Wilson, N. H. M. (2006). Potential uses of transit smart card registration and transaction data to improve transit planning. *Transportation Research Record: Journal of the Transportation Research Board*(1971), 119-126. doi: 10.3141/1971-16
- Venner, M. (2013). *GPS/AVL use, challenges, and cost-benefit in operations*. Communication présentée à Transportation Research Board 92nd Annual Meeting, Washington, DC. Tiré de <http://amonline.trb.org/pap@PaperNo=13-5123>
- Ville de Montréal. (2008). Plan de transport 2008. Tiré de <http://servicesenligne.ville.montreal.qc.ca/sel/publications/htdocs/porteacespublicationFr/porteacespublication.jsp?systemName=68235660>
- Ville de Montréal. (2010). Bilan 2010 du Plan de transport 2008. Tiré de <http://servicesenligne.ville.montreal.qc.ca/sel/publications/htdocs/porteacespublicationFr/porteacespublication.jsp?systemName=94985911>
- Volosin, S. E., Paul, S., Pendyala, R. M., Grady, B., & Gardner, B. (2012). *The Application of a microsimulation model system to the analysis of a light-rail corridor: insights from a TRANSIMS deployment*. Communication présentée à Transportation Research Board 91st Annual Meeting, Washington, DC. Tiré de <http://pubsindex.trb.org/view/2012/C/1130944>
- Voyer, P. (2000). *Tableaux de bord de gestion et indicateurs de performance* (2^e éd.). Québec, QC: Presses de l'Université du Québec.

- Vuchic, V. R. (2005). *Urban transit: operations, planning and economics*. Hoboken, NJ: John Wiley & Sons.
- Vuchic, V. R. (2007). *Urban transit: systems and technology*. Hoboken, NJ: John Wiley & Sons.
- Wang, Y., Zhu, Y., He, Z., Yue, Y., & Li, Q. (2011). *Challenges and opportunities in exploiting large-scale GPS probe data* (Rapport n° HPL-2011-109). Hewlett Packard Laboratories. Tiré de <http://www.hpl.hp.com/techreports/2011/HPL-2011-109.html>
- Wilson, N. H. M., & Nuzzolo, A. (2009). *Schedule-based modeling of transportation networks: theory and applications*. New York, NY: Springer.
- Wilson, N. H. M., Zhao, J., & Rahbee, A. (2009). The potential impact of automated data collection systems on urban public transport planning. Dans Nigel H.M. Wilson and Agostino Nuzzolo (Édit.), *Schedule-based modeling of transportation networks: theory and applications* (p. 75-99). New York, NY: Springer.
- Yoh, A. C., Iseki, H., Taylor, B. D., & King, D. A. (2006). Interoperable transit smart card systems: are we moving too slowly or too quickly? *Transportation Research Record: Journal of the Transportation Research Board*(1986), 69-77. doi: 10.3141/1986-11
- yWorks. (2013). yED Graph Editor (Version 3.10.2) [Logiciel]. Tiré de http://www.yworks.com/en/products_yed_about.html
- Zhao, J. (2004). *The planning and analysis implications of automated data collection systems: rail transit OD matrix inference and path choice modeling examples*. (Mémoire de maîtrise, Massachusetts Institute of Technology, Cambridge, MA). Tiré de <http://dspace.mit.edu/handle/1721.1/28752>
- Zhao, J., Rahbee, A., & Wilson, N. H. M. (2007). Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. *Computer-Aided Civil and Infrastructure Engineering*, 22(5), 376-387. doi: 10.1111/j.1467-8667.2007.00494.x

ANNEXES

ANNEXE 1 Informations sur certaines figures

Les figures suivantes utilisent des icônes trouvés à partir de la banque d'images IconFinder (<http://www.iconfinder.com>). Pour alléger les figures, les informations sur la provenance des icônes sont mises dans ce tableau :

figure	icône	référence
2-2	autobus	<i>Farm-fresh</i> par FatCow Web Hosting, 2013. Tiré de http://www.iconfinder.com/icondetails/92551/32/bus_icon
	tours à bureaux	<i>Mixed 3d icons</i> par IFA, 2013. Tiré de http://www.iconfinder.com/icondetails/104633/256/buildings_enterprise_icon
	personne qui marche	<i>IconSweets 2</i> par Yummygum, 2013. Tiré de http://www.iconfinder.com/icondetails/63071/40/user_walk_icon
	maison	<i>Vista Inspirate</i> par Alexandre Moore, 2013. Tiré de http://www.iconfinder.com/icondetails/2860/32/home_house_icon
	voiture	<i>Vista Style Transport</i> par Icons Land, 2013. Tiré de http://www.iconfinder.com/icondetails/22990/64/car_cars_grey_jawahar_transportation_vehicle_icon
	dollar	<i>Farm-fresh</i> par FatCow Web Hosting, 2013. Tiré de http://www.iconfinder.com/icondetails/92551/32/bus_icon
	personnes	icônes par défaut de yED Graph Editor, par yWorks (2013)
3-5	satellites	<i>Sophistique Night</i> par Iconshock, 2013. Tiré de http://www.iconfinder.com/icondetails/45140/128/satellite_icon
	autobus	(création personnelle)
	ordinateur de bord	(création personnelle)
	autres icônes	icônes par défaut de yED Graph Editor, par yWorks (2013)
2-4	autobus	<i>Farm-fresh</i> par FatCow Web Hosting, 2013. Tiré de http://www.iconfinder.com/icondetails/92551/32/bus_icon
annexe 5	serveur et personne	icônes par défaut de yED Graph Editor, par yWorks (2013)

ANNEXE 2 Définitions d'objets

Cette annexe résume la définition de plusieurs objets mentionnés dans ce mémoire. Ces définitions sont issues et inspirées du Transit Capacity and Quality of Service Manual (Transportation Research Board & Kittelson and Associates Incorporated, 2003) et du glossaire du Public Utility Research Center (2006).

objet	définition
arrêt	lieu où les usagers peuvent attendre, monter, descendre et correspondre entre les véhicules
autobus	véhicule autopropulsé conçu pour accueillir un grand nombre de passagers — habituellement plus de 16 usagers — et circulant sur le réseau routier
battement (temps de ~)	temps entre deux voyages productifs successifs Cela peut inclure les temps de repos et de recouvrement.
carte à puce	support de plastique muni d'une puce électronique où sont stockés les titres de transport de son détenteur
correspondance	changement de véhicule effectué par un usager lors d'un déplacement
déplacement	mouvement spatiotemporel d'une personne entre deux points, une origine et une destination, dans un seul sens et pour un motif spécifique
descendant, débarquant*	usager qui sort d'un véhicule
descente, débarquement*	action, fait de sortir d'un véhicule
éjection (voyage d'~, <i>pull-in</i>)	HLP entre le dernier arrêt du dernier voyage productif d'une tournée et le garage
garage, dépôt	lieu de stationnement et de maintenance des autobus
halte	action, fait d'immobiliser le véhicule
haut-le-pied (HLP)	voyage — improductif — s'effectuant sans passagers
injection (voyage d'~, <i>pull-out</i>)	HLP entre le garage et le premier arrêt du premier voyage productif d'une tournée
interligne	HLP entre deux voyages productifs Le temps alloué à l'interligne peut inclure les temps de repos et de recouvrement.
ligne	service de transport regroupant une combinaison de parcours desservant approximativement les mêmes endroits Le regroupement est souvent déterminé arbitrairement par l'autorité de transport et est identifié par un numéro.
ligne-tracé	parcours desservant une séquence précise d'arrêts et pouvant être décrit par son numéro de ligne associée par sa direction
ligne-direction	ensemble des lignes-tracés d'une même ligne qui vont dans la même direction

(La suite du tableau est à la page suivante.)

ANNEXE 2 Définition d'objets (suite)

objet	définition
montant, embarquant*	usager qui entre dans un véhicule
montée, embarquement*	action, fait d'entrer dans un véhicule
point d'éjection (~ du réseau)	dernier arrêt du dernier voyage productif d'une tournée, avant que l'autobus ne retourne au garage
point d'injection (~ dans le réseau)	arrêt où un autobus débute son premier voyage productif
point de contrôle	arrêt ou lieu où le temps de passage du véhicule est vérifié par rapport à l'horaire planifié
point de montée, de descente	arrêt où l'usager est entré ou sorti du véhicule
recouvrement, récupération (temps de ~)	temps prévu après un voyage productif au cas où un retard se serait accumulé pour réduire au minimum les effets sur le voyage productif suivant
remise en disponibilité, rotation (temps de ~)	temps nécessaire à un véhicule pour effectuer un aller-retour sur une ligne, incluant le temps de battement
repos (temps de ~)	temps prévu pour la pause du chauffeur après un voyage productif
réseau	ensemble relié de nœuds et de liens système opérationnellement coordonné de lignes d'autobus
terminal, gare	ensemble aménagé d'infrastructures, d'installations et d'équipements pouvant accueillir un grand nombre d'usagers et servant de point important de départ ou de fin ou de correspondance à des lignes de transport
terminus	point d'arrêt extrême d'un voyage
terminus de départ	premier arrêt d'un voyage
terminus d'arrivée	dernier arrêt d'un voyage
ticket, passage	titre de transport donnant accès au réseau sur une ligne une seule fois Dans le cas de la STM, le passage est accompagné d'un droit de correspondance (voir conditions à la section 3.3.1).
titre de transport	droit de passage, limité ou non et payé d'avance, sur le réseau
tournée	séquence de voyages affectée à un autobus entre sa sortie du garage et son retour
trajet, itinéraire, parcours	chemin spécifique pour un autobus
validation (~ de passage)	action d'acquiescement de son droit de passage par argent comptant ou par un titre de transport valide, à l'aide de la boîte de perception de l'autobus
voyage (<i>trip</i>)	mouvement d'un autobus dans une direction à un moment donnée entre deux points spécifiques (terminus, garage, point d'injection, etc.)
voyage productif, commercial	voyage au cours duquel l'autobus peut transporter des usagers

* « Embarquement » et « débarquement » et leurs dérivés sont réservés aux navires, aux embarcations, aux avions ou aux trains et sont à éviter pour les véhicules routiers (OQLF, 2002).

ANNEXE 3 Estimation des véhicules-heures des voyages GTFS associés à un voyage réel

heure	véhicules-heures	% local étendu	% local	% express	% local pointe	% express pointe	% navette	% nuit	% battement	% interligne	% HLP garage début	% HLP garage fin
4	7.5	25.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	73.2	0.0
5	48.3	60.5	0.0	0.0	0.0	0.0	0.0	0.0	4.2	5.8	29.6	0.0
6	101.5	56.3	1.1	5.9	2.9	2.3	0.0	0.0	6.7	8.0	16.7	0.0
7	160.8	53.9	3.9	9.9	1.7	7.2	0.0	0.0	6.5	6.8	9.1	1.1
8	162.4	53.8	5.0	7.9	2.5	6.1	0.0	0.0	6.2	10.0	0.2	8.4
9	118.5	55.8	5.0	8.1	0.7	6.2	0.2	0.0	8.7	8.4	0.0	6.9
10	99.6	64.7	4.1	9.1	0.0	1.1	0.0	0.0	12.6	7.5	0.0	0.9
11	96.3	64.1	4.6	10.3	0.0	0.0	0.0	0.0	11.8	6.1	1.0	2.1
12	94.8	67.8	3.1	9.8	0.0	0.0	0.0	0.0	12.3	5.1	1.0	0.8
13	96.1	70.5	2.5	9.2	0.0	0.0	0.0	0.0	12.1	5.3	0.0	0.3
14	106.3	62.8	2.8	10.2	0.2	0.0	0.0	0.0	8.7	4.6	10.6	0.2
15	157.0	51.2	2.1	8.9	2.7	3.7	0.0	0.0	6.9	2.0	21.6	0.9
16	157.6	54.2	2.6	7.8	1.7	9.6	0.0	0.0	7.3	9.5	5.9	1.5
17	158.3	56.6	3.5	6.0	1.7	9.0	0.0	0.0	8.1	11.8	0.3	3.1
18	134.8	60.6	2.5	3.9	0.7	7.0	0.0	0.0	8.9	8.2	0.0	8.1
19	96.0	68.6	1.6	2.2	0.0	1.1	0.0	0.0	10.1	7.7	0.0	8.7
20	71.6	74.4	1.1	1.0	0.0	0.0	0.0	0.0	9.7	10.9	0.0	2.9
21	66.1	73.9	1.2	0.1	0.0	0.0	0.0	0.0	8.6	11.1	0.5	4.6
22	53.4	74.8	0.0	0.0	0.0	0.0	0.0	0.0	10.4	10.6	0.0	4.2
23	45.3	78.9	0.0	0.0	0.0	0.0	0.0	0.0	9.8	6.2	0.0	5.1
24	36.8	76.2	0.0	0.0	0.0	0.0	0.3	1.4	10.3	6.4	0.0	5.5
25	22.6	45.6	0.0	0.0	0.0	0.0	0.0	20.6	7.4	5.3	0.0	21.1
26	6.6	0.3	0.0	0.0	0.0	0.0	0.0	67.3	10.4	6.6	0.0	15.5
27	5.7	0.0	0.0	0.0	0.0	0.0	0.0	74.3	18.4	5.0	0.0	2.3
28	4.6	0.0	0.0	0.0	0.0	0.0	0.0	82.4	11.9	5.4	0.0	0.4
29	2.4	0.0	0.0	0.0	0.0	0.0	0.0	45.5	0.0	0.0	0.0	54.5
Total	2110.8	60.3	2.7	6.5	1.0	3.7	0.0	0.9	8.7	7.6	5.2	3.5

ANNEXE 4 Variations AFC-APC par arrêt (véhicule 22289, ligne 139)

nom d'arrêt	tracé 1 (vers le sud)						tracé 2 (vers le nord)							
	n° séq	nombre	centiles (AFC - APC)					n° séq	nombre	centiles (AFC - APC)				
			0	25	50	75	100			0	25	50	75	100
Notre-Dame	1	3	-2	-1	0	0	0	44	-	-	-	-	-	-
Jeanne-d'Arc / Sainte-Catherine	2	5	-1	0	0	0	0							
Sainte-Catherine	3	6	-2	-2	-1	0	4	43	9	-1	0	0	0	0
Adam	4	7	-4	-1	0	0	1	42	7	0	0	0	0	0
La Fontaine	5	5	-1	0	0	0	1	41	7	0	0	0	0	0
Ontario	6	8	-13	-2	0	0	0	40	10	0	0	0	0	0
Voie ferrée C.N.	7	6	-1	0	0	0	2	39	7	0	0	0	0	0
de Rouen	8	7	-2	-2	-1	0	0	38	7	-1	0	0	0	0
Hochelaga	9	5	-2	0	0	0	1	37	9	-2	0	0	0	1
Pierre-de-Coubertin	10	9	-40	-19	-1	0	3	36	10	-33	-3	-2	0	0
Sherbrooke	11	7	0	0	0	0	0							
Rachel	12	4	0	0	0	0	0	35	9	-1	0	0	0	0
du Mont-Royal	13	7	0	0	0	0	0	34	6	-1	0	0	0	0
Saint-Joseph	14	7	0	0	0	0	1	33	10	-4	0	0	0	1
Laurier	15	6	-1	0	0	0	0	32	9	-1	0	0	0	0
Masson	16	8	0	0	0	0	1	31	9	-1	0	0	0	0
Dandurand	17	6	0	0	0	0	0	30	7	0	0	0	1	1
Rosemont	18	8	0	0	0	0	5	29	10	-2	0	0	0	2
Bellechasse	19	7	0	0	0	0	1	28	8	0	0	0	1	1
Beaubien	20	8	-5	-1	0	0	2	27	10	-8	0	0	0	1
Saint-Zotique	21	6	0	0	0	0	1	26	8	0	0	0	0	0
Bélanger	22	9	-1	0	0	0	1	25	9	-2	-1	0	0	1
Bélair	23	7	0	0	0	0	0	24	4	0	0	0	0	0
Jean-Talon	24	8	-4	-1	0	0	1	23	10	-7	-1	0	1	2
Éverett	25	4	0	0	0	0	0	22	5	-1	0	0	0	1
Villeray	26	7	-3	0	0	0	0	21	8	-4	-1	0	0	1
Jarry	27	9	-3	-1	0	0	1	20	10	-13	-2	0	0	0
d'Hérelle	28	5	0	0	0	0	0	19	7	-1	0	0	0	0
Jean-Rivard	29	8	0	0	0	0	1	18	10	-1	0	0	0	0
Robert	30	8	0	0	0	0	1	17	10	-1	0	0	0	1
Denis-Papin	31	6	-1	0	0	0	0	16	8	-1	-1	0	0	1
39e Rue	32	6	0	0	0	0	0	15	10	-3	-2	-1	0	1
43e Rue	33	5	0	0	0	0	0	14	7	-2	-1	0	0	0
45e Rue	34	4	0	0	0	0	0							
47e Rue								13	10	-2	0	0	1	1
48e Rue	35	9	0	0	0	0	1							
des Grandes-Prairies	36	9	-1	0	0	0	0	12	10	-1	0	0	1	1
55e Rue / 56e rue	37	5	0	0	0	0	0	11	8	0	0	0	0	1
Industriel	38	4	0	0	0	0	0	10	3	0	0	0	0	0
Majeau / Larin	39	6	0	0	0	0	0	9	5	0	0	0	0	1
Fleury								8	8	0	0	0	0	1
Forest	40	6	-1	0	0	0	0	7	10	0	0	0	0	1
Prieur								6	9	-1	0	0	0	0
d'Amiens	41	8	-1	0	0	0	0							
Monselet								5	9	0	0	0	0	1
de Castille	42	8	0	0	0	0	1							
Charlevoix	43	8	0	0	0	0	0	4	7	-7	0	0	0	0
Arthur-Champoux	44	2	0	0	0	0	0	3	7	-8	-1	0	0	0
Gariépy / Henri-Bourassa								2	9	-27	-6	0	0	0
d'Amos / Saint-Julien	45	5	0	0	0	0	0							
d'Amos / Garon	46	5	0	0	0	0	0							
Garon / Henri-Bourassa	47	8	-12	0	0	0	0							
Henri-Bourassa / Pie-IX	48	-	-	-	-	-	-	1	9	-14	-3	-1	0	0

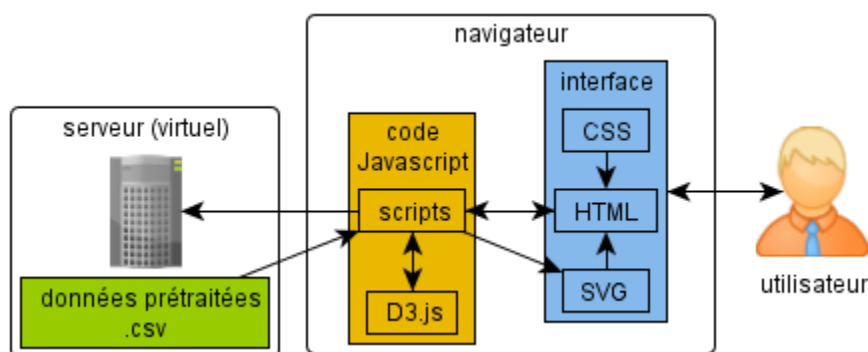
*n°Séq : numéro de l'arrêt dans la séquence / nombre : nombre de passages détectés

ANNEXE 5 Détails sur les outils interactifs (diagrammes distance-temps et tableaux de bord)

Le tableau ci-dessous montre les versions des standards Web sur lesquelles s'appuient les outils interactifs et les navigateurs compatibles. La compatibilité est limitée pour les navigateurs dont la version est marquée d'une étoile (*) : une boîte de texte avec le code hexadécimal de la couleur est affichée au lieu du bouton de choix de couleur (voir annexe 7 et annexe 9). Le code fait appel à un serveur, au moins virtuel ou local comme WampServer (Bourdon & Leclerc, 2013), pour appeler les fichiers de données prétraitées sous format tabulaire (.csv). En se basant sur les données prétraitées, le code permet de dessiner les différents éléments du diagramme en format vectoriel SVG. Le code HTML inclut des éléments d'interface pour que l'utilisateur puisse modifier les éléments des dessins.

catégorie	élément	version	notes
standards Web	HTML	5	
	CSS	3	
	SVG	1.1	
bibliothèque JS	D3.js	3	compatible avec la version 2 aussi
navigateurs compatibles testés	Google Chrome	28	
	Opera	12.16	
	Firefox	22	*
	Internet Explorer	10	*

La figure ci-dessous résume le fonctionnement simplifié du code JavaScript et de ses relations avec les différents éléments en jeu :



(Voir annexe 1 pour références des icônes)

ANNEXE 6 Fichiers entrants pour le diagramme distance-temps

Exemple avec le corridor Saint-Michel

Fichier « arrêts » :

(distance en mètres) (« bullet » : arrêt pour service local et express ou arrêt « important »)

id	name	dist	bullet
0	Gouin (67→S) / Henri-Bourassa (467→S)	-1040	1
1	Gouin (67→N)	-700	1
2	Henri-Bourassa (67/467→N)	-634	1
3	Prieur	-279	0
4	Fleury	0	1
5	Mont-Joli	352	0
6	Sauvé	534	1
7	Charland	855	0
8	Champdoré	1085	0
9	de Louvain	1317	1
10	Legendre	1598	0
11	Émile-Journault	1878	1
12	Denis-Papin	2161	0
13	Robert	2441	1
14	Déville	2725	0
15	d'Hérelle	3003	0
16	d'Hérelle (67→N)	3090	0
17	Jarry	3277	1
18	Crémazie	3519	0
19	Crémazie (67→S)	3607	0
20	Villeray	3832	1
21	L.-O.-David	4051	0
22	(↓) MÉTRO Saint-Michel	4248	1
23	Jean-Talon	4344	0
24	Bélair	4620	0
25	Bélanger	4744	1
26	Saint-Zotique	5085	0
27	Beaubien	5414	1
28	Bellechasse	5751	0
29	Rosemont	5927	1
30	Holt	6096	0
31	Dandurand	6294	0
32	Masson	6568	1
33	Laurier	6809	0
34	Saint-Joseph	6977	1
35	du Mont-Royal	7284	0
36	William-Tremblay	7535	0
37	Rachel / Davidson	7779	0
38	Davidson / Sherbrooke (→N)	7997	0
39	Rachel / de Chambly (→S)	8070	0
40	Hochelaga / Davidson (→N)	8284	0
41	Valois / Sherbrooke (→S)	8404	1
42	Nicolet / Hochelaga (→S)	8908	0
43	(↓) MÉTRO Joliette (→N)	8549	1
44	(↓) MÉTRO Joliette (→S)	9040	1

Fichier « passages » (extrait) :

trip	ligne	trace	arretID	temps	voyageReel	arrivee	depart	charge	montees	descentes	veh
029_002_002	67	1	43	22380	14128665	22293	22372	9	9	0	26085
029_002_002	67	1	40	22440	14128665	22415	22424	9	0	0	26085
029_002_002	67	1	38	22500	14128665	22460	22475	10	1	0	26085
029_002_002	67	1	37	22560	14128665	22514	22535	10	1	1	26085
029_002_002	67	1	36	22560	14128665	22561	22582	10	0	0	26085
029_002_002	67	1	35	22620	14128665	22609	22609	10	0	0	26085
029_002_002	67	1	34	22680	14128665	22646	22658	11	2	1	26085
029_002_002	67	1	33	22740	14128665	22679	22679	11	0	0	26085
029_002_002	67	1	32	22800	14128665	22706	22725	11	1	1	26085
029_002_002	67	1	31	22860	14128665	22758	22767	12	2	1	26085
029_002_002	67	1	30	22920	14128665	22796	22799	13	1	0	26085
029_002_002	67	1	29	22920	14128665	22825	22879	15	3	1	26085
029_002_002	67	1	28	22980	14128665	22904	22940	15	1	1	26085
029_002_002	67	1	27	23040	14128665	22976	22984	19	4	0	26085
029_002_002	67	1	26	23160	14128665	23014	23014	19	0	0	26085
029_002_002	67	1	25	23220	14128665	23045	23080	21	2	0	26085
029_002_002	67	1	24	23220	14128665	23094	23094	21	0	0	26085
...

Fichier « lignes-tracés-directions » :

```

var service = {
  "67": {
    "type": "R",
    "couleurGTFS": "rgb(204,192,218)",
    "couleurReels": "rgb(128,100,162)",
    "couleurCharges": "rgb(201,193,255)",
    "2": "S",
    "1": "N",
    "3": "N",
  },
  "467": {
    "type": "E",
    "couleurGTFS": "rgb(141,180,226)",
    "couleurReels": "rgb(54,96,146)",
    "couleurCharges": "rgb(133,202,255)",
    "1": "S",
    "2": "N",
  },
};

//Fichiers
var fichierStations = "line_patterns_10J_S_DET_67_467.csv", //arrêts
    fichierDonnees = "donnees_67_467.csv"; //passages

```

Diagramme illustrant la structure du fichier « lignes-tracés-directions » :

- Le service est défini par un objet `service` contenant deux lignes : "67" et "467".
- La ligne "67" est de type "R" (local) et la ligne "467" est de type "E" (express).
- Les tracés et directions sont définies par les clés "1", "2" et "3" pour la ligne "67", et "1" et "2" pour la ligne "467".
- Les couleurs GTFS, Reels et Charges sont définies pour chaque ligne.
- Les fichiers de données sont définis : `fichierStations` pour les arrêts et `fichierDonnees` pour les passages.

ANNEXE 7 Paramétrage du diagramme distance-temps, lignes 67 et 467

affichage et propriétés
des objets du diagramme

dimensions du diagramme
et heures à afficher

ligne	voyages		charge	direction		MD	
	GTFS	« réels »		N ↑	S ↓	montées	descentes
locale	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
express	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

☐ voyages GTFS réalisés seulement ☒ légende des charges et des MD

grille		texte	titres	
arrêts express	<input type="checkbox"/>	1 px	arrêts	11 px
arrêts local	<input type="checkbox"/>	1 px	heures	11 px
heures	<input type="checkbox"/>	1 px	légende	11 px

texte

police

taille px

charges

épaisseur min px

échelle pers./px

opacité %

MD

échelle px²/pers.

opacité %

options du diagramme

largeur* px

hauteur* px

heures

heure min+ h

heure max+ h

espaces pour étiquettes

paddingX px

paddingY px

* espaces pour étiquettes non inclus dans largeur et hauteur

+ heures depuis minuit

arrêts

Gouin (67→S) / Henri-Bourassa (467→S)

Henri-Bourassa (67/467→N)

charges / MD

10

Prieur

Fleury

heures depuis minuit → 8

9

ANNEXE 8 Fichier entrant pour le tableau de bord

Extrait du fichier entrant montrant la ligne 467, tracé 1 (direction sud)

ligne	tracé	stopSeq	stopID	nom	terminus	nboist	centdist000	centdist009	centdist050	centdist091	centdist100	nboistcumu1	centdistcumu000	centdistcumu009	centdistcumu050	centdistcumu091	centdistcumu100	nboistarrivee	centPoncArrivee000
467	1	1	1802186	Saint-Michel / Henri-Bourassa	0	35	0	0	0	0	0	35	0	0	0	0	0	35	-73
467	1	2	1802198	Saint-Michel / Henri-Bourassa															
467	1	3	1800223	Saint-Michel / Fleury	0							35	613	1061	1075	1105	1162	35	-131
467	1	4	105683	Saint-Michel / Sauvé	0	35	495	522	525	535	541	35	1146	1586	1601	1630	1687	35	-15
467	1	5	106783	Saint-Michel / de Louvain	0	35	743	763	765	769	786	35	1932	2345	2365	2398	2455	35	-17
467	1	6	107613	Saint-Michel / Émile-Journault	0	35	520	541	552	555	568	35	2499	2895	2917	2948	3007	35	-17
467	1	7	108413	Saint-Michel / Robert	0	35	531	547	551	563	616	35	3066	3449	3467	3531	3557	35	-15
467	1	8	109593	Saint-Michel / Jarry	0	35	742	803	822	826	846	35	3912	4259	4285	4325	4379	35	-23
467	1	9	110323	Saint-Michel / Villeray	0	35	514	531	541	556	572	35	4471	4803	4825	4869	4923	35	-17
467	1	10	110913	Station Saint-Michel	0	35	374	393	409	415	420	35	4891	5208	5236	5275	5331	35	-19
467	1	11	111763	Saint-Michel / Bélanger	0	35	465	481	484	504	577	35	5391	5694	5723	5784	5814	35	-10
467	1	12	112973	Saint-Michel / Beaubien	0	35	568	644	656	668	693	35	6066	6346	6379	6439	6471	35	-11
467	1	13	114103	Saint-Michel / Rosemont	0	35	473	501	503	515	516	35	6582	6860	6881	6923	6975	35	-14
467	1	14	115583	Saint-Michel / Masson	0	35	594	622	625	635	669	35	7226	7476	7508	7560	7600	35	-15
467	1	15	116673	Saint-Michel / Saint-Joseph	0	35	357	388	401	413	433	35	7639	7874	7910	7952	7998	35	-16
467	1	16	119473	Valois / Sherbrooke	0	35	1255	1393	1409	1425	1442	35	9081	9282	9316	9371	9407	35	-21
467	1	17	142352	Station de métro Joliette	1	35	542	554	629	662	752	35	9744	9869	9929	10011	10035	35	-22

ANNEXE 9 Paramétrage du tableau de bord, ligne 467 sud

choix de la ligne-tracé

propriétés des objets

taille des colonnes

Ligne-tracé : 467 1 → Station de métro Joliette

BAM	boîtes	lignes	portes	1	2	3	axes	texte
distance			AFC				général	
arrivée			montées				avance	
départ			descentes				à l'heure	
durée			AFC - APC				retard	
<input checked="" type="checkbox"/> Toujours conserver BAM au premier arrêt et au terminus pour la ponctualité			polylignes charges 1 px cumul AFC 1 px cumul M 1 px			(AFC-APC) O 1 px charge O 1 px		

titre des axes

11 px

texte

police Trebuchet MS

taille données 11 px

taille verticale

hauteur* 1340 px

paddingY 60 px

blocs

marge+ 11 px

arrêts 190 px

nb 15 px

dist. 35 px

dist. BAM 5 px

ponct. 200 px

durée 85 px

blocs (suite)

AFC-APC 50 px

AFC 75 px

montées 75 px

descentes 75 px

charge 75 px

cumul M 75 px

* espace pour les titres des axes et les étiquettes non inclus

+ entre chaque bloc

- nombre de passages

Les boîtes à moustaches (BAM) montrent les centiles 0, 9, 50, 91 et 100.

Actualiser

arrêts	nb	dist.	avance (s) retard →	durée halte (s)	AFC - APC	validations	montées APC	descentes APC	cumul charge	cumul montants
Saint-Michel / Henri-Bourassa__1	35	0	-240 -60 0 180	540 0 60 120 180	-50 240	10 140 280 420 560 700	0 140 280 420 560 700	0 140 280 420 560 700	0 1290	3110